PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/10, C07K 1/14 — A2

(11) International Publication Number: **WO 97/05248**

(43) International Publication Date: 13 February 1997 (13.02.97)

(21) International Application Number: PCT/US96/11875

(22) International Filing Date: 18 July 1996 (18.07.96)

(30) Priority Data:
08/509,164 31 July 1995 (31.07.95) US

(71)(72) Applicant and Inventor: CHOMCZYNSKI, Piotr [US/US]; 778 Avon Fields Lane, Cincinnati, OH 45229 (US).

(74) Agents: AHRENS, Gregory, F. et al.; Wood, Herron & Evans, P.L.L., 2700 Carew Tower, Cincinnati, OH 45202 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: SOLUTION CONTAINING CHAOTROPIC AGENT AND PROCESS USING IT FOR ISOLATION OF DNA, RNA AND PROTEINS

(57) Abstract

Solutions and methods are disclosed for the effective, simple isolation/extraction of DNA, RNA and proteins from a single biological material sample, such as cells, tissues and biological fluids. The preferred solutions include effective amounts of a chaotropic agent(s), buffer, reducing agent, and may or may not include an organic solvent. Genomic DNA and total RNA can be isolated utilizing the solutions and methods of the invention in as little as 20 minutes, and proteins in as little as 30 minutes.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

SOLUTION CONTAINING CHAOTROPIC AGENT AND PROCESS USING IT FOR ISOLATION OF DNA, RNA AND PROTEINS

Field of the Invention

The invention relates to compositions and methods for isolating nucleic acids (both RNA and DNA) and proteins from biological materials. And more particularly, the invention relates to nucleic acid and protein isolation methods employing non-toxic chaotropic agents.

Background of the Invention

The continuous advances in molecular biology, biotechnology and clinical research have resulted in an ever increasing number of uses for DNA, RNA and proteins. For example, polymerase chain reaction (PCR) technology has dramatically expanded the use of DNA and RNA in basic research, in clinical diagnostics such as detection of messenger RNA (mRNA) by reverse transcription PCR (RT-PCR), and the use of PCR in detection of genetic defects. In the protein field, identification of proteins by Western blotting has become an important tool in studying gene expression in basic research and identification of specific proteins for diagnostic purposes, as exemplified by viral protein detection.

The increased use of RNA, DNA and proteins has created a need for fast, simple and reliable methods and reagents for isolating DNA, RNA and proteins. In many applications, collecting the biological material sample and subsequent analysis thereof would be substantially simplified if the three cellular components (RNA, DNA and proteins) could be simultaneously isolated from a single sample. The simultaneous isolation is especially important when the sample size is so small, such as in biopsy, that it precludes its separation into smaller samples to perform separate isolation protocols for DNA, RNA and proteins.

There are known methods for isolating DNA, RNA and proteins from a single biological material sample. One such method is described in Coombs, L. M., et al.: "Simultaneous Isolation of DNA, RNA and Antigenic Protein Exhibiting Kinase Activity from Small Tumor Samples Using Guanidine Isothiocyanate", Anal. Biochem., 188, 338-343 (1990). The Coombs et al. method is based on ultracentrifugation of the sample homogenate in a guanidine-cesium chloride solution. The sample is homogenized in 4 M guanidine thiocyanate and then overlayered on a cesium chloride (CsCl) solution and centrifuged at >100,000g for about 18 hours. Following centrifugation, DNA, RNA and proteins are separated and purified over the next 12-24 hours. This method has several limitations or drawbacks, including the prolonged time required for isolation and the limited number and size of samples which can be processed with an ultracentrifuge. Also, the high cost of an ultracentrifuge may be prohibitive in certain circumstances.

Another method for the simultaneous isolation of DNA, RNA and proteins from a single biological material sample is the subject of my earlier U.S. Patent No. 5,346,994 (the '994 patent). That method is based on liquid-phase separation using phenol and guanidine thiocyanate. A biological sample is homogenized in the aqueous solution of phenol and guanidine thiocyanate and the homogenate thereafter is mixed with chloroform. Following centrifugation, the homogenate separates into an organic phase, an interphase and an aqueous phase. Proteins are sequestrated into the organic phase, DNA into the interphase and RNA into the aqueous phase. Next, each component is precipitated from the corresponding phase using ethanol and is then washed. The whole procedure can be completed in about 2-3 hours, and is especially useful for the isolation of high quality RNA from a variety of sources. One drawback to this method is the use of highly toxic phenol.

There are many known methods for the separate isolation of DNA, RNA and proteins from biological material; i.e., protocols for isolating a single one of these components from a sample. In typical DNA isolation methods, a biological sample is lysed in a lysing solution and then the DNA is isolated from the lysate according to any one of a variety of multi-step protocols, which may take from one hour to several days to complete. Frequently recommended DNA isolation methods involve the use of toxic phenol. See, Sambrook, J. et al., "Molecular Cloning", Vol. 2, pp. 9.14-9.23, Cold Spring Harbor Laboratory Press (1989) and Ausubel, Frederick M. et al., "Current Protocols in Molecular Biology", Vol. 1, pp. 2.2.1-2.4.5, John Wiley & Sons, Inc. (1994). Typically, a biological sample is lysed in a detergent

solution and the protein component of the lysate is digested with proteinase for 12 - 18 hours. Next, the lysate is extracted with phenol to remove most of the cellular components, and the remaining aqueous phase is processed further to isolate DNA. In another method, described in Van Ness et al. U.S. Patent 5,130,423, non-corrosive phenol derivatives are used for the isolation of nucleic acids. The resulting preparation is a mix of RNA and DNA.

DNA isolation methods utilizing non-corrosive chaotropic agents have also been developed. These methods, which are based on the use of guanidine salts, urea and sodium iodide, involve lysis of a biological sample in a chaotropic aqueous solution and subsequent precipitation of the crude DNA fraction with a lower alcohol. The final purification of the precipitated, crude DNA fraction can be achieved by any one of several methods, including column chromatography as described in "RapidPrep™ Genomic DNA Isolation Kits For Cells and Tissue: Versatility at Your Fingertips!", Analects, Vol 22, No. 4, Pharmacia Biotech, 1994, or exposure of the crude DNA to a polyanion-containing protein as described in Koller U.S. Patent 5,128,247.

Yet another method of DNA isolation. which is described in Botwell, D.D.L., "Rapid Isolation of Eukaryotic DNA", Anal. Biochem. 162, 463-465 (1987) involves lysing cells in 6 M guanidine hydrochloride, precipitating DNA from the lysate at acid pH by adding 2.5 volumes of ethanol, and washing the DNA with ethanol. It is believed that the resulting DNA may be contaminated, however, with a low molecular weight material such as RNA and pigments. This conclusion is in agreement with the well known report showing that under similar conditions RNA can be precipitated from cell or

tissue lysate. See Chirgwin, J.M. et al., "Isolation of Biologically Active Ribonucleic Acid from Sources Enriched in Ribonuclease", Dept. of Biochemistry and Biophysics, Univ. of California, Vol. 18, No. 24, pp. 5294-5299 (1979).

Sodium iodide, another chaotropic agent, has been used in DNA isolation, but its use requires an additional purification step consisting of adsorption of the isolated DNA on glass beads. Although this method is relatively simple, it results in a low yield of isolated DNA.

In still another approach, the bulk of cytoplasmic proteins and RNA are removed by lysing cell samples in a detergent solution. The lysate is then fractionated into the nuclear and cytoplasmic fractions. And thereafter, the DNA is purified by dissolving the nuclear fraction in a chaotropic solution, precipitating and washing with ethanol. This method, described in Ciulla, T.A. et al., "A Simple Method for DNA Purification from Peripheral Blood", Anal. Biochem. 174, 485-488 (1988), can be completed in about 2 hours and is useful for isolating DNA from whole blood.

Known techniques for isolating RNA typically utilize either guanidine salts or phenol extraction, as described for example in Sambrook, J. et al., "Molecular Cloning", Vol. 1, pp. 7.3-7.24, Cold Spring Harbor Laboratory Press (1989) and Ausubel, F. M. et al., "Current Protocols in Molecular Biology", Vol. 1, pp. 4.0.3-4.4.7, John Wiley & Sons, Inc. (1994). Phenol-based techniques are multi-step procedures requiring several hours or days to complete. Similarly, the guanidine-based RNA isolation methods require at least several hours and take many steps to complete. In my earlier

U.S. Patent No. 4,843,155, phenol and guanidine procedures were uniquely combined, resulting in a simple method of total RNA isolation which can be completed in 3 hours. The method of the '155 patent was further improved upon in my '994 patent, which allows for completion of the RNA isolation in about 1 hour.

There are also known techniques for the simultaneous isolation of DNA and RNA, as referenced in my earlier '994 patent, the disclosures of which are incorporated herein by reference. All of these techniques utilize phenol extraction as a necessary step for isolating RNA and DNA free of protein contamination.

Heretofore, it has been the commonly accepted view that precipitation of nucleic acids from chaotropic solvents does not result in pure nucleic acid preparations. Contrary to this view, however, the present inventor has found that under certain conditions, as described in full detail herein, the use of chaotropic agents alone will result in isolation of assay ready, high quality DNA and RNA. This unexpected finding led to the development of a very simple, effective method and product for the simultaneous isolation of DNA, RNA and proteins from a single sample for subsequent use in molecular biology, biotechnology, clinical research and other applications.

Summary of the Invention

The products and methods of the present invention provide a highly effective, simple means of extracting DNA, RNA and proteins from a single biological material sample, such as cells, tissues and biological fluids. Advantageously, these results can be achieved without the use of toxic or

corrosive reagents and without the use of expensive ultracentrifugation equipment. Genomic DNA and total RNA can be isolated utilizing the products and methods of the invention in as little as 20 minutes, and proteins in as little as 30 minutes. These results are substantially faster than existing methods for the simultaneous isolation of DNA, RNA and proteins. The invention is also applicable to the separate isolation of RNA only, DNA only, proteins only or any combination of two of these cellular components. The resulting genomic DNA and total RNA isolated utilizing the methods and products of the invention are of high quality suitable for use in research, biotechnology, etc. The invention is in part based on the unexpected finding that utilizing the products of the invention, RNA is precipitated prior to DNA, which is contrary to prior art methodologies. In particular, the finding that the solvent solution of the present invention precipitates RNA prior to DNA is in striking contrast to the described prior art such as Koller U.S. Patent No. 5,128,247, wherein it describes that DNA exhibits a lower solubility than RNA and apparently can be precipitated more easily than RNA. Furthermore, substantially lower amounts of organic solvents are required to effect the precipitation of the cellular components.

In its broadest aspects, the invention encompasses solutions for isolating substantially pure and undegraded RNA, DNA and protein from biological materials, including tissue, cells and fluids. The solution preferably includes effective amounts of a chaotropic agent(s), buffer, reducing agent, and water (with or without organic solvent(s)). The chaotropic agent(s) act to dissociate proteins from nucleic acids (RNA and DNA) and inhibit the activity

of nucleic acid degradation enzymes. The actions of the chaotropes are potentiated by the reducing agent present in the solvent solution. When the solution includes an organic solvent, its presence prevents solubilization of RNA, thereby making it possible to remove RNA from the homogenate formed when the tissue sample is homogenized in the solvent solution by a brief centrifugation step. When the solution does not include an organic solvent, the RNA is precipitated from the homogenate subsequent to homogenization by the addition of an organic solvent.

Preferred chaotropic agents for the solution include guanidine thiocyanate, guanidine hydrochloride, and mixtures thereof. These components may be supplemented with other chaotropes, such as urea or sodium iodide. The preferred concentration of chaotropes in the solution is in the range of about 2 M - 7 M. Preferably, the reducing agent is non-toxic, such as 2-aminoethanethiol. If required, this can be substituted with 2-mercaptoethanol; however, this is a toxic composition. The reducing agent facilitates denaturization of RNase by the chaotropes and aids in the isolation of undegraded RNA.

The solution of the present invention preferably contains a sufficient amount of buffer to maintain the pH of the solution above about 6. For the simultaneous isolation of RNA, DNA and proteins, the pH should be maintained in the range of about 6-7.5. For the isolation of DNA alone, the effective pH range may be about 6-12. The preferred buffers for use in the solutions of the invention include tris-hydrochloric acid, sodium phosphate, sodium acetate, sodium tetraborate-boric acid and glycine-sodium hydroxide.

As stated, the solution may contain an organic solvent. The preferred solvents are lower alcohols such as methanol, ethanol and isopropanol. However, other water miscible solvents can be used which will achieve the desired effect, such as acetone, polyethylene glycol and dimethylsulfoxide, or mixtures of any of the foregoing. The effective concentration of organic solvents in the product of the invention is in the range of about 15-30% by volume.

The solution of the invention may contain additional components, including organic and inorganic salts such as sodium chloride, potassium chloride, ammonium chloride, sodium phosphate, sodium acetate, sodium nitrite, lithium chloride, and sodium bromide. Furthermore, compatible detergents such as sarcosines and polyoxyethylenesorbitan, and chelating agents such as ethylenediamine tetraacetic acid and citric acid can be utilized to promote tissue solubilization and precipitation of nucleic acids.

In another aspect, the invention encompasses methods of isolating substantially pure RNA, DNA and proteins from biological material samples. Utilizing the solutions of the present invention, and particularly those wherein the solution includes an organic solvent, the method includes an initial step of homogenizing a biological material sample in the solution to form an homogenate. The presence of the organic solvent prevents solubilization of RNA, thereby making it possible to remove RNA from the homogenate by brief centrifugation (sedimentation). Thereafter, DNA is precipitated from the remaining homogenate by adding an effective amount of an organic solvent and recovering the precipitated DNA by brief centrifugation (sedimentation).

Finally, adding an effective amount of an organic solvent to the remaining homogenate precipitates proteins therefrom. The successive addition of organic solvents to precipitate genomic DNA and proteins from the post-RNA homogenate require a limited amount of the organic solvent. Following a wash with ethanol, the three cellular components are each separately and fully ready for use in molecular biology, biotechnology and clinical research applications. Preferably, the organic solvent utilized to precipitate DNA is added in a ratio of about 0.15-0.3 volumes of solvent per one volume of initial homogenate. The organic solvent added to precipitate proteins is preferably added in the ratio of about 3-4 volumes of solvent per one volume of initial homogenate. One preferred organic solvent for this purpose is isopropanol, although other suitable organic solvents can be used.

In an alternative methodology, wherein the solution does not include an organic solvent component, the homogenate is formed and an initial centrifugation step is performed to remove any unhomogenized tissue. Thereafter an effective amount of an organic solvent is added to preclude the solubilization of total RNA. This makes it possible to recover the RNA from the homogenate by a brief centrifugation (sedimentation). Preferably, the organic solvent added to precipitate RNA is added in a ratio of about 0.15-0.3 volumes of solvent per one volume of initial homogenate. Once separated, the RNA is washed with ethanol and the RNA precipitate may be dissolved in formamide and stored at -20°C. The use of formamide as a solubilization agent is beneficial in that it protects RNA from degradation by RNase, which may otherwise contaminate the isolated RNA.

Following precipitation of the RNA, DNA is precipitated by adding an additional 0.15-0.3 volumes of organic solvent per one volume of the initial homogenate. The DNA precipitate is removed from the homogenate by spooling or brief centrifugation (sedimentation). Proteins and other cellular components are retained in the homogenate and the protein component is precipitated by the addition of 3-4 volumes of an organic solvent per one volume of the initial homogenate.

Overall, and in comparison with other known methodologies, the precipitation of nucleic acids according to the present invention is performed with a substantially reduced amount of organic solvents. For example, in Sambrook, J. et al., "Molecular Cloning", Vol. 1, pp. 7.3-7.24, Cold Spring Harbor Press (1989), the precipitation of RNA requires the addition of 0.5-2.5 volumes of a lower alcohol added to one volume of the chaotropic solution. In addition, the RNA precipitation was carried out for several hours at -20-4°C. In contrast, the precipitation of RNA according to the present method is completed in about 3-5 minutes and is performed at room temperature. With respect to the DNA precipitation, Botwell, D.D.L. "Rapid Isolation of Eukaryotic DNA", Anal. Biochem. 162, 463-465 (1987), describes precipitation from chaotropic solutions carried out with 2-2.5 volumes of ethanol. Likewise, high volumes of alcohol have been recommended for the effective DNA precipitation from non-chaotropic solutions, as exemplified in Ausubel, F.M. et al., "Current Protocols in Molecular Biology", Vol. 1, pp. 221-245, John Wiley & Sons, Inc. (1994).

The unexpectedly low concentration of organic solvents required for precipitating RNA and DNA from the chaotropic solutions of the present invention makes it feasible to obtain RNA and DNA in highly pure form for use in molecular biology and biotechnology, as well as clinical research and other applications.

These and other features and advantages of the present invention will become apparent to persons skilled in the art upon review of the detailed description and working examples herein.

Brief Description of the Drawings

Fig. 1A shows the electrophoretic analysis of nucleic acids isolated in accordance with Example 1 herein;

Fig. 1B shows the utility of the total RNA in RT-PCR;

Fig. 1C shows the utility of the genomic DNA in PCR;

Fig. 1D shows the results of protein analysis by Western blotting;

Fig. 2A shows the electrophoretic analysis of nucleic acids isolated in accordance with Example 2 herein;

Fig. 2B shows the results of total RNA analysis by Northern blotting;

Fig. 2C shows the results of genomic DNA analysis by Southern blotting;

Fig. 2D shows the results of protein analysis by Western blotting; and

Fig. 3 shows the results of genomic DNA analysis by electrophoresis as isolated in Example 5.

Detailed Description of the Invention

Preferred solutions and methods according to the present invention are described in the following working Examples.

Example 1 - Simultaneous Isolation of DNA, RNA & Proteins From Cells

A method of the invention was used to simultaneously isolate DNA, RNA and proteins from rat somatomammotroph P0 cells. About 10^8 P0 cells were lysed (homogenized) in a 10ml of solution of the invention, which contained: 4 M guanidine thiocyanate (Amresco, Inc., Solon, OH), 17% isopropanol, 0.1 M sodium acetate, 0.1 M 2-aminoethanethiol hydrochloride (Sigma, St. Louis, MO) and 0.2% sarkosyl in water. The solution was adjusted to pH 7.0 with hydrochloric acid. Unless stated otherwise, chemical reagents were obtained from Fisher Scientific (Pittsburgh, PA). Next, 0.6 ml aliquots of the lysate (homogenate) were frozen or used immediately for the isolation. The 0.6 ml aliquot utilized as described below contained 37.3 μg DNA, as determined by the diphenylamine reaction.

RNA isolation. The lysate (0.6 ml) was centrifuged for 8 minutes at 10,000 g at room temperature to sediment total RNA. The post-RNA lysate was transferred to a fresh tube and saved for the DNA and protein isolation described below. The RNA pellet was washed with 1 ml of 95% ethanol by vortexing and pipetting off the ethanol. Finally, RNA was dissolved in formamide and stored at -20 C. The total RNA isolation was completed in 11 minutes. The isolated RNA exhibited a 260/280 ratio of 1.79 $\pm$ 0.05, with a yield of 49.2 $\pm$ 2.8 μg RNA (mean $\pm$ SD, n=3). Northern blotting of the

isolated RNA preparations showed an undegraded pattern of mRNA when tested for growth hormone, prolactin, β-actin and GAPDH mRNAs.

For use in the reverse transcription PCR (RT-PCR), an aliquot of the RNA-solubilized formamide was precipitated with 3 volumes of ethanol. The precipitate was dissolved in water and used for RT-PCR.

It has been found that spectrophotometrical measurements to determine the optical density of RNA and DNA are substantially improved when performed in a solution containing higher than usual concentrations of chelating agents. For this purpose, the concentration of chelating agent(s) should be higher than 5 mM, with the optimum for ethylenediamine tetraacetate (EDTA) at 10 mM and for citrate at 30 mM. This is a new and unexpected finding. Typically, optical density readings of RNA and DNA are performed in water or 1 mM EDTA.

At higher concentration of chelating agents, the optical density readings are more reproducible and result in a higher 260/280 ratio. For example, the RNA preparation described in Example 1 had a 260/280 ratio of 1.79 ± 0.05 when measured in water, and a ratio of 1.97 ± 0.01 when measured in 10 mM EDTA.

DNA isolation. DNA was precipitated from the post-RNA lysate by the addition of 0.15 ml of isopropanol. The floating DNA precipitate was swirled (spooled) onto the pipette tip and transferred to a new tube. The remaining post-DNA lysate was saved for the protein isolation described below. The DNA was washed by mixing it with 1 ml of 95% ethanol and pipetting out the ethanol wash. The final DNA preparation was dissolved in 8 mM NaOH by

gentle pipetting followed by neutralization of the solution with N-[2-hydroxyethyl] piperazine-N-[2-ethane sulfinic acid] (HEPES, free acid). The isolation was completed in less than 14 minutes. The isolated DNA exhibited a 260/280 ratio 1.81 ± 0.02 (SD, n=3), indicating a lack protein contamination. The average yield from three isolations was 33.9 ± 2.9 (SD, n=3) μg DNA. As compared with the original amount of DNA in the lysate, determined by the diphenylamine reaction, the method of the invention provided 91% recovery of DNA.

Protein isolation. Proteins were precipitated from the post-DNA lysate by the addition of 1.8 ml of isopropanol and centrifugation at 10,000 g for 5 minutes. The precipitate was washed with 95% ethanol and dissolved in 0.1 N acetic acid. Alternatively, the precipitate could be dissolved in 0.1% sodium dodecylsulphate, or in water. The protein isolation was completed in 22 minutes. The isolated protein preparation was tested by Western blotting using a specific anti-rat prolactin antibody. The presence of the prolactin specific band in Western blotting is indicative of the good quality of the isolated protein preparation.

The results of tests performed with the simultaneously isolated RNA, DNA and proteins are shown in Figure 1. Figure 1 A shows the results of nucleic acids electrophoresed in 1% agarose gel and stained with ethidium bromide. Lanes 1 and 2 show undegraded total RNA (3 μg/lane); lanes 3 and 4 show high molecular weight genomic DNA (3 μg/lane); and lanes 5 and 6 show genomic DNA (3 μg) digested for 2 hours with EcoR1 restrictase. As the results demonstrate, there is no detectable DNA in the RNA preparation and no

detectable RNA in the DNA preparation. The total digestion of DNA by EcoR1 restrictase is indicative of the good quality of the isolated DNA. Figures 1B-D show the results of RT-PCR, PCR and Western blotting performed with total RNA, genomic DNA and proteins, respectively, isolated as described hereinabove. The RT-PCR (Fig. 1B) and PCR (Fig. 1C) were performed using glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and rat growth hormone (GH) primers, respectively, and Western blotting (Fig. 1C) was performed using anti-rat prolactin antibody. Amplification of the 374 base pair GAPDH DNA fragment in RT-PCR and the 686 base pair GH DNA fragment in PCR indicates that the isolated nucleic acid preparations (RNA and DNA) are adequately purified for the PCR reaction. Also, the presence of a prolactin specific band in Western blotting shows the high quality of the isolated protein preparation.

Similar results in the simultaneous isolation of RNA, DNA and proteins have been obtained when, in place of isopropanol, other water miscible organic solvents such as ethanol, methanol, acetone, dimethylsulfoxide, polyethylene glycol or mixtures of these solvents are used. These water miscible organic solvents can be used as components of the lysing solution and/or as the precipitating agents. All solvents are available from companies such Aldrich Chemical Co., Inc. (Milwaukee, WI) or Fluka Chemical Corp. (Ronkonkoma, NY).

EXAMPLE 2 - Simultaneous Isolation of RNA, DNA and Proteins From Tissues.

In this embodiment of the method, the solution does not include an organic solvent component. This permits the removal of any unhomogenized

tissue fragments from the lysate (homogenate) by a brief initial centrifugation. Thereafter, RNA is precipitated from the clear lysate by the addition of 0.2-0.3 volumes of an organic solvent.

A frozen sample of rat pituitary was homogenized in a hand held glass-Teflon homogenizer with 1 ml of a lysing solution containing 4 M guanidine thiocyanate, 0.1 M sodium acetate, 0.1 M 2-aminoethanethiol hydrochloride and 0.2% sarkosyl in water. The solution was adjusted to pH 7.0 with hydrochloric acid. The homogenate was centrifuged at 10,000 g for 5 minutes. The clear supernatant was transferred to a new tube, mixed with 0.3 ml of isopropanol and stored for 3 minutes at room temperature to precipitate RNA. The precipitated RNA was removed by centrifugation at 10,000 g for 8 minutes and dissolved in formamide. The post-RNA supernatant was further processed to isolate DNA and proteins in the same manner as described in Example 1. DNA and proteins were successively precipitated from the post-RNA lysate by the addition of 0.5 ml acetone and 10.75 ml of acetone, respectively. The isolated RNA exhibited a 260/280 ratio of 1.74 and the yield was 0.06 mg. The DNA exhibited a 260/280 ratio of 1.78 and the yield was 0.04 mg. As shown in Figure 2A, there was no detectable DNA present in the RNA preparation and no detectable RNA present in the DNA preparation. As in Fig. 1A, the nucleic acids were electrophoresed in 1% agarose gel and stained with ethidium bromide. Lanes 1 and 2, total RNA (3 μg/lane); lanes 3 and 4, genomic DNA (3 μg/lane); and lanes 5 and 6, genomic DNA (3 μg) digested for 2 hours with EcoR1 restrictase.

The preparations of total RNA, genomic DNA and proteins were tested by Northern, Southern and Western blotting for the rat growth hormone (GH) mRNA, GH gene, and GH, respectively. All three analyses, depicted in Figs. 2B-D, respectively, show that the method of invention yielded high quality RNA, DNA and protein preparations.

EXAMPLE 3. Isolation of RNA From Cells without Centrifugation of The Initial Lysate

Breast tumor MCF7 cells were grown in monolayer culture in a 3.5 cm petri dish. At the end of the culture period, the culture medium was removed and the cells were lysed by adding 1 ml of the lysing solution directly to the culture dish. The lysing solution used was that in Example 1. The lysate was centrifuged at 10,000 g for 5 minutes, and the RNA pellet was washed with 95% ethanol and dissolved in formamide. The isolated RNA exhibited a 260/280 ratio 1.81 and the yield was 22 μg RNA.

EXAMPLE 4. Isolation of RNA from Tissue with Centrifugation of The Initial Lysate to Remove Tissue Fragments

Rat kidney (0.95 g) was homogenized in 19 ml of a lysing solution having the following composition: 4 M guanidine thiocyanate, 0.1 M sodium acetate and 0.2% sarkosyl in water. The solution was adjusted to pH 7.0 with hydrochloric acid. The homogenate was centrifuged at 10,000 g for 5 minutes to remove unsolubilized material. The resulting supernatant was transferred to a new tube and RNA was precipitated from the supernatant by the addition of 3.8 ml (0.3 volume) of isopropanol and centrifugated at 10,000 g for 8 minutes. The RNA pellet was dissolved in formamide and stored at -20 C.

The RNA preparation exhibited a 260/280 ratio 1.77 and the yield was 3.9 mg of RNA.

EXAMPLE 5 - Isolation of DNA from Cells without Centrifugation

The lysing solution used for this DNA isolation contained the following: 4 M guanidine thiocyanate, 0.1 M sodium acetate, 17% isopropanol, 0.2% sarkosyl in water. The lysing solution was adjusted to pH 9 by the addition of 0.4 M NaOH.

Rat pituitary P0 cells were lysed in the lysing solution by repetitive pipetting. The lysate was mixed with 0.4 volume of ethanol and the precipitated DNA was spooled onto a pipette tip and washed twice with 95% ethanol. The resulting DNA was dissolved in 8 mM NaOH and neutralized to pH 8.0 with 0.1 M HEPES buffer.

This embodiment of the invention further shortens the DNA isolation protocol by omitting centrifugation of the lysate. This deletion of centrifugation results in only a minor contamination of the isolated DNA with RNA. Analysis of the genomic DNA isolated as described above was performed by electrophoreses thereof in 1% agarose gel and staining with ethidium bromide. Lanes 1 and 2 of Fig. 3 are genomic DNA (3 μg/lane), and lanes 3 and 4 are genomic DNA digested for 2 hours with EcoR1. As is apparent in Figure 3, only a residual amount of RNA was detected in the low molecular weight region of the agarose gel by the ethidium bromide staining. Apparently, spooling DNA onto a pipette tip removes mainly DNA, while most of the partially hydrolyzed RNA (as indicated by its low molecular weight) remains in the lysate. The presence of a reducing agent is not necessary for the isolation

of DNA and the lysing solution can have pH within the range 6 - 12. The alkaline pH can be adjusted with NaOH, KOH or other organic or inorganic alkaline reagents. The best results are believed to be obtained when the lysate has a pH between 8 - 9.

The unexpected and rapid hydrolysis of RNA occurring at pH 8 - pH 9 can be attributed to the presence of the chaotropic agent. This new finding allows for isolation of good quality DNA in a simple, one-step method.

The protocol for DNA isolation without centrifugation can be completed in less than 10 minutes. This is believed to be the fastest and simplest method of genomic DNA isolation. Importantly, the only equipment required for this method are tubes and pipettes. This allows for performing the DNA isolation on a field trip or elsewhere with a limited access to laboratory equipment.

EXAMPLE 6 - Isolation of DNA from Tissues with One Centrifugation of the Lysate

Rat spleen (127 mg) was homogenized in a hand held glass-Teflon homogenizer with 5 ml of the lysing solution described in Example 5. The homogenate was centrifuged at 10,000 g for 5 minutes. The supernatant was transferred to a new tube and DNA was precipitated and washed as in Example 5. The isolated DNA exhibited a 260/280 ratio 1.76, and the yield was 1.93 mg DNA.

The scope of the present invention is not intended to be limited to the specific Examples described herein, but is to be accorded a scope commensurate with the appended claims.

What is claimed is:

1. A solution for isolating substantially pure and undegraded RNA, DNA and proteins from biological material, said solution comprising:

effective amounts of

at least one chaotropic agent, and

a buffer present in an amount sufficient to maintain the pH of said solution in the range of about 6-12.

2. The solution of claim 1 wherein said chaotropic agent is selected from the group consisting of guanidine thiocyanate, guanidine hydrochloride, urea, sodium iodide and mixtures thereof.

3. The solution of claim 2 wherein said chaotropic agent is present at a concentration in the range of about 2M-7M.

4. The solution of claim 1 wherein said buffer is selected from the group consisting of tris-hydrochloric acid, sodium phosphate, sodium acetate, sodium citrate, sodium tetraborate and glycine-sodium hydroxide.

5. The solution of claim 1 wherein said buffer is present in an amount sufficient to maintain the pH of said solution in the range of about 6-7.5.

6. The solution of claim 1 further comprising at least one reducing agent.

7. The solution of claim 6 wherein said reducing agent is selected from the group consisting of 2-aminoethanethiol, 2-mercaptoethanol and mixtures thereof.

8. The solution of claim 1 further comprising a salt.

9. The solution of claim 8 wherein said salt is selected from the group consisting of sodium chloride, potassium chloride, ammonium chloride, sodium acetate, sodium nitrate, lithium chloride, sodium bromide and mixtures thereof.

10. The solution of claim 1 further comprising at least one chelating agent.

11. The solution of claim 10 wherein said chelating agent is selected from the group consisting of ethylenediamine tetraacetic acid and citric acid.

12. The solution of claim 1 further comprising a detergent.

13. The solution of claim 12 wherein said detergent is selected from the group consisting of sarcosine and polyoxyethylenesorbitan.

14. A solution for isolating substantially pure and undegraded RNA, DNA and proteins from biological material, said solution comprising:

effective amounts of

at least one chaotropic agent,

a buffer present in an amount sufficient to maintain the pH of said solution in the range of 6-12, and

an organic solvent present in an amount comprising about 15-30% of said solution.

15. The solution of claim 14 wherein said organic solvent is selected from the group consisting of lower alcohols, acetone, polyethylene glycol, dimethylsulfoxide, and mixtures thereof.

16. The solution of claim 14 wherein said chaotropic agent is selected from the group consisting of guanidine thiocyanate, guanidine hydrochloride, urea, sodium iodide, and mixtures thereof.

17. The solution of claim 16 wherein said chaotropic agent is present at a concentration in the range of about 2M-7M.

18. The solution of claim 14 wherein said buffer is selected from the group consisting of tris-hydrochloric acid, sodium phosphate, sodium acetate, sodium citrate, sodium tetraborate and glycine-sodium hydroxide.

19. The solution of claim 14 wherein said buffer is present in an amount sufficient to maintain the pH of said solution in the range of about 6-7.5.

20. The solution of claim 14 further comprising at least one reducing agent.

21. The solution of claim 20 wherein said reducing agent is selected from the group consisting of 2-aminoethanethiol, 2-mercaptoethanol and mixtures thereof.

22. The solution of claim 14 further comprising a salt.

23. The solution of claim 22 wherein said salt is selected from the group consisting of sodium chloride, potassium chloride, ammonium chloride, sodium acetate, sodium nitrate, lithium chloride, sodium bromide and mixtures thereof.

24. The solution of claim 14 further comprising at least one chelating agent.

25. The solution of claim 24 wherein said chelating agent is selected from the group consisting of ethylenediamine tetraacetic acid and citric acid.

26. The solution of claim 14 further comprising a detergent.

27. The solution of claim 26 wherein said detergent is selected from the group consisting of sarcosine and polyoxyethylenesorbitan.

28. A method of isolating substantially pure and undegraded RNA, DNA and proteins from biological material, comprising the steps of:

a) homogenizing a biological material sample in a solution of claim 1 to form an homogenate;

b) removing unhomogenized material from the homogenate by sedimentation;

c) precipitating RNA in the remaining homogenate by adding an effective amount of an organic solvent thereto, and recovering the precipitated RNA by sedimentation;

d) thereafter precipitating DNA in the remaining homogenate by adding an effective amount of organic solvent thereto, and recovering the precipitated DNA by one of spooling and sedimentation; and

e) thereafter precipitating proteins from the remaining homogenate by adding an effective amount of an organic solvent thereto, and recovering the precipitated proteins by sedimentation.

29. The method of claim 28 wherein said organic solvent added to precipitate RNA is added in a ratio of about 0.15-0.3 volumes of solvent per one volume of initial homogenate.

30. The method of claim 29 wherein said organic solvent is selected from the group consisting of lower alcohols, acetone, polyethylene glycol and dimethylsulfoxide.

31. The method of claim 28 wherein said organic solvent added to precipitate DNA is added in a ratio of about 0.15-0.3 volumes of solvent per one volume of initial homogenate.

32. The method of claim 31 wherein said organic solvent is selected from the group consisting of lower alcohols, acetone, polyethylene glycol and dimethylsulfoxide.

33. The method of claim 28 wherein said organic solvent added to precipitate proteins is added in a ratio of about 3-4 volumes of solvent per one volume of initial homogenate.

34. The method of claim 33 wherein said organic solvent is selected from the group consisting of lower alcohols, acetone, polyethylene glycol and dimethylsulfoxide.

35. A method of isolating substantially pure and undegraded RNA, DNA and proteins from biological material, comprising the steps of:

a) homogenizing a biological material sample in the solution of claim 14 to form an homogenate;

b) recovering substantially pure, undegraded RNA from the homogenate by sedimentation;

c) thereafter precipitating DNA in the remaining homogenate by adding an effective amount of an organic solvent thereto, and recovering the precipitated DNA by one of sedimentation and spooling; and

d) thereafter precipitating proteins from the remaining homogenate by adding an effective amount of organic solvent thereto, and recovering the precipitated proteins by sedimentation.

36. The method of claim 35 wherein said organic solvent added to precipitate DNA is added in a ratio of 0.15-0.3 volumes of solvent per one volume of the initial homogenate.

37. The method of claim 36 wherein said organic solvent is selected from the group consisting of lower alcohols, acetone, polyethylene glycol and dimethylsulfoxide.

38. The method of claim 35 wherein said organic solvent added to precipitate proteins is added in a ratio of 3-4 volumes of solvent per one volume of the initial homogenate.

39. The method of claim 38 wherein said organic solvent is selected from the group consisting of lower alcohols, acetone, polyethylene glycol and dimethylsulfoxide.

40. A method of isolating substantially pure and undegraded RNA from biological material, comprising the steps of:

a) homogenizing a biological material sample in a solution of claim 1 to form an homogenate;

b) removing unhomogenized material from the homogenate by sedimentation; and

c) precipitating RNA in the homogenate by adding an effective amount of an organic solvent thereto, and recovering the precipitated RNA by sedimentation.

41. The method of claim 40 wherein said organic solvent is selected from the group consisting of lower alcohols, acetone, polyethylene glycol and dimethylsulfoxide.

42. A method of isolating substantially pure and undegraded RNA from biological material, comprising the steps of:

a) homogenizing a biological material sample in a solution of claim 14 to form an homogenate; and

b) recovering substantially pure RNA from the homogenate by sedimentation.

43. A method of isolating substantially pure and undegraded DNA from biological material, comprising the steps of:

a) homogenizing a biological material sample in a solution of claim 14 to form an homogenate; and

b) precipitating DNA in the homogenate by adding an effective amount of an organic solvent thereto, and recovering the precipitated DNA by one of spooling and sedimentation.

44. The method of claim 43 wherein said organic solvent is selected from the group consisting of lower alcohols, acetone, polyethylene glycol and dimethylsulfoxide.

45. A method of isolating substantially pure and undegraded DNA from biological material, comprising the steps of:

a) homogenizing a biological material in the solution of claim 1 to form an homogenate;

b) removing unhomogenized material from the homogenate by sedimentation; and

c) precipitating DNA in the homogenate by adding an effective amount of an organic solvent thereto, and recovering the precipitated DNA by one of spooling and sedimentation.

46. The method of claim 45 wherein said organic solvent is selected from the group consisting of lower alcohols, acetone, polyethylene glycol and dimethylsulfoxide.

FIG. 1A

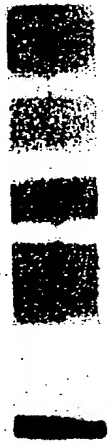

FIG. 1B

FIG. 1C

FIG. 1D

FIG. 2A

FIG. 2B

FIG. 2C

FIG. 2D

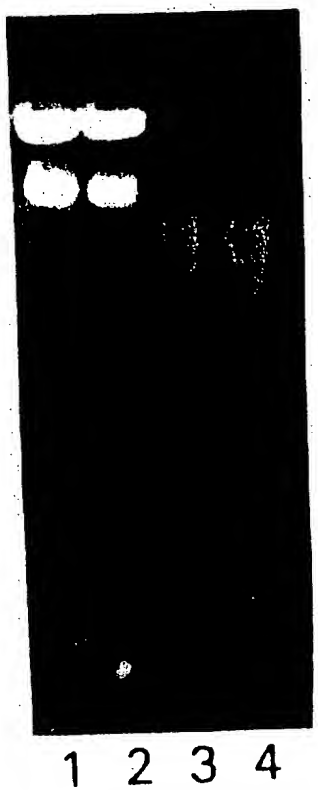

FIG. 3

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/10, C07K 1/14 | A3

(11) International Publication Number: **WO 97/05248**

(43) International Publication Date: 13 February 1997 (13.02.97)

(21) International Application Number: PCT/US96/11875

(22) International Filing Date: 18 July 1996 (18.07.96)

(30) Priority Data:
08/509,164 31 July 1995 (31.07.95) US

(71)(72) Applicant and Inventor: CHOMCZYNSKI, Piotr [US/US]; 778 Avon Fields Lane, Cincinnati, OH 45229 (US).

(74) Agents: AHRENS, Gregory, F. et al.; Wood, Herron & Evans, P.L.L., 2700 Carew Tower, Cincinnati, OH 45202 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, UZ, VN, ARIPO patent (KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
6 March 1997 (06.03.97)

(54) Title: SOLUTION CONTAINING CHAOTROPIC AGENT AND PROCESS USING IT FOR ISOLATION OF DNA, RNA AND PROTEINS

(57) Abstract

Solutions and methods are disclosed for the effective, simple isolation/extraction of DNA, RNA and proteins from a single biological material sample, such as cells, tissues and biological fluids. The preferred solutions include effective amounts of a chaotropic agent(s), buffer, reducing agent, and may or may not include an organic solvent. Genomic DNA and total RNA can be isolated utilizing the solutions and methods of the invention in as little as 20 minutes, and proteins in as little as 30 minutes.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

INTERNATIONAL SEARCH REPORT

Int onal Application No
PCT/US 96/11875

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/10 C07K1/14

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US,A,5 128 247 (KOLLER CHARLES A) 7 July 1992 see column 9 - column 10 ---	1-27, 40-46
X	ANALYTICAL BIOCHEMISTRY, vol. 207, no. 1, 15 November 1992, US, pages 197-201, XP000611804 NELSON J.E. AND KRAWETZ S.A.: "Purification of cloned and genomic DNA by guanidine thiocyanate/isobutyl alcohol fractionation" see abstract see page 198, column 1 --- -/--	1-27, 43-46

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
12 December 1996	29. 01 97

Name and mailing address of the ISA	Authorized officer
European Patent ffice, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax (+31-70) 340-3016	Chakravarty, A

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Intr onal Application No
PCT/US 96/11875

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	J. MICROBIOL. METHODS, vol. 19, no. 3, 1994, pages 167-172, XP000610898 LEMA M.W ET AL.: "A general method for the extraction of DNA from bacteria" see page 168 ---	1-27,45, 46
X	EP,A,0 544 034 (SIEMENS AG) 2 June 1993 see the whole document ---	1-5,8,9
X	US,A,5 162 507 (WOLFE SIDNEY N ET AL) 10 November 1992 see column 7, line 48 - line 68 ---	1-11
X	WO,A,92 00983 (MICROPROBE CORP) 23 January 1992 see examples 1,4 -----	1-5

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int ional Application No
PCT/US 96/11875

Patent document cited in search report	Publication date	Patent family member(s)		Publication date
US-A-5128247	07-07-92	AU-A-	6432490	03-04-91
		CA-A-	2064068	15-02-91
		EP-A-	0487646	03-06-92
		WO-A-	9102740	07-03-91
EP-A-0544034	02-06-93	CA-A-	2083870	29-05-93
		JP-A-	5300174	12-11-93
US-A-5162507	10-11-92	AU-B-	622860	30-04-92
		AU-A-	1627988	06-12-88
		EP-A-	0368857	23-05-90
		JP-T-	2503384	18-10-90
		NO-B-	176797	20-02-95
		WO-A-	8808849	17-11-88
		CA-A-	1337671	28-11-95
WO-A-9200983	23-01-92	CA-A-	2087105	14-01-92
		EP-A-	0539515	05-05-93
		US-A-	5393672	28-02-95
		US-A-	5130423	14-07-92

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/68 — A1

(11) International Publication Number: WO 97/35034

(43) International Publication Date: 25 September 1997 (25.09.97)

(21) International Application Number: PCT/US97/04269

(22) International Filing Date: 19 March 1997 (19.03.97)

(30) Priority Data:
60/013,835 20 March 1996 (20.03.96) US

(60) Parent Application or Grant
(63) Related by Continuation
US 60/013,835 (CON)
Filed on 20 March 1996 (20.03.96)

(71) Applicant *(for all designated States except US)*: THE GOVERNMENT OF THE UNITED STATES OF AMERICA, represented by THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES [US/US]; Bethesda, MD 20892 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: GONZALEZ, Frank, J. [US/US]; 5000 Battery Lane, Unit 101, Bethesda, MD 20814 (US). FERNANDEZ-SALGUERO, Pedro [ES/US]; 4863 Battery Lane #22, Bethesda, MD 20814 (US).

(74) Agents: ALICEA, Hector, A. et al.; Townsend and Townsend and Crew L.L.P., 8th floor, Two Embarcadero Center, San Francisco, CA 94111 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: METHODS AND COMPOSITIONS FOR DETECTING DIHYDROPYRIMIDINE DEHYDROGENASE SPLICING MUTATIONS

(57) Abstract

The present invention provides compositions, methods and kits for the detection of genetic polymorphisms or mutations related to dihydropyrimidine dehydrogenase deficiency (DPDD). The polymorphism or mutations generally occur in the dihydropyrimidine dehydrogenase DPD gene in chromosome 1. Also provided are mutant forms of DPD.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

METHODS AND COMPOSITIONS FOR DETECTING DIHYDROPYRIMIDINE DEHYDROGENASE SPLICING MUTATIONS

TECHNICAL FIELD OF THE INVENTION

The present invention relates to methods and compositions for detecting splicing defects in the dihydropyrimidine dehydrogenase gene. The methods and compositions are useful for identifying persons who are at risk of a toxic reaction to the commonly employed cancer chemotherapy agent, 5-fluorouracil.

BACKGROUND OF THE INVENTION

5-Fluorouracil (5-FU) is commonly used in the treatment of cancers, including cancers of the breast, head, neck, and digestive system. The efficacy of 5-FU as a cancer treatment varies significantly among patients. Clinically significant differences in systemic clearance and systemic exposure of 5-FU are often observed. *See*, Grem, J.L. *In* Chabner, B.A. and J.M. Collins (eds.), *Cancer Chemotherapy: Principles and Practice*, pp. 180-224, Philadelphia, PA, Lippincott, 1990). Furthermore, 5-FU treatment is severely toxic to some patients, and has even caused death. *See*, Fleming *et al.* (1993) *Eur. J. Cancer* 29A: 740-744; Thyss *et al.* (1986) *Cancer Chemother. Pharmacol.* 16: 64-66; Santini *et al.* (1989) *Br. J. Cancer* 59: 287-290; Goldberg *et al.* (1988) *Br. J. Cancer* 57: 186-189; Trump *et al.* (1991) *J. Clin. Oncol.* 9: 2027-2035; and Au *et al.* (1982) *Cancer Res.* 42: 2930-2937.

Patients in whom 5-FU is severely toxic typically have low levels of dihydropyrimidine dehydrogenase (DPD) activity. *See*, Tuchman *et al.* (1985) *N. Engl. J. Med.* 313: 245-249; Diasio *et al.* (1988) *J. Clin. Invest.* 81: 47-51; Fleming *et al.* (1991) *Proc. Am. Assoc. Cancer Res.* 32: 179; Harris *et al.* (1991) *Cancer (Phila.)* 68: 499-501; Houyau *et al.* (1993) *J. Nat'l. Cancer Inst.* 85: 1602-1603; Lyss *et al.* (1993) *Cancer Invest.* 11: 239-240. Dihydropyrimidine dehydrogenase (DPD, EC 1.3.1.2) is the principal enzyme involved in the degradation of 5-FU, which acts by inhibiting thymidylate synthase. *See*, Heggie *et al.* (1987) *Cancer Res.* 47: 2203-2206; Chabner *et al.* (1989) *In* DeVita *et al.* (eds.), *Cancer - Principles and Practice of Oncology*, pp. 349-

395, Philadelphia, PA, Lippincott; Diasio *et al.* (1989) *Clin. Pharmacokinet.* 16: 215-237; Grem *et al.*, *supra*. The level of DPD activity also affects the fficacy of 5-FU treatments, as 5-FU plasma levels are inversely correlated with the level of DPD activity. *See*, Iigo *et al.* (1988) *Biochem. Pharm.* 37: 1609-1613; Goldberg *et al.*, *supra.*; Harris *et al.*, *supra.*; Fleming *et al.*, *supra*. In turn, the efficacy of 5-FU treatment of cancer is correlated with plasma levels of 5-FU.

In addition to its 5-FU degrading activity, DPD is also the initial and rate limiting enzyme in the three-step pathway of uracil and thymine catabolism, leading to the formation of ß-alanine and β-aminobutyric acid, respectively. *See*, Wasternack *et al.* (1980) *Pharm. Ther.* 8: 629-665. DPD deficiency is associated with inherited disorders of pyrimidine metabolism, clinically termed thymine-uraciluria. *See*, Bakkeren *et al.* (1984) *Clin. Chim. Acta.* 140: 247-256. Clinical symptoms of DPD deficiency include a nonspecific cerebral dysfunction, and DPD deficiency is associated with psychomotor retardation, convulsions, and epileptic conditions. *See*, Berger *et al.* (1984) *Clin. Chim. Acta* 141: 227-234; Wadman *et al.* (1985) *Adv. Exp. Med. Biol.* 165A: 109-114; Wilcken *et al.* (1985) *J. Inherit. Metab. Dis.* 8 (Suppl. 2): 115-116; van Gennip *et al.* (1989) *Adv. Exp. Med. Biol.* 253A: 111-118; Brockstedt *et al.* (1990) *J. Inherit. Metab. Dis.* 12: 121-124; and Duran *et al.* (1991) *J. Inherit. Metab. Dis.* 14: 367-370. Biochemically, patients having DPD deficiency have an almost complete absence of DPD activity in fibroblasts (*see*, Bakkeren *et al.*, *supra*) and in lymphocytes (*see*, Berger *et al.*, *supra* and Piper *et al.* (1980) *Biochim. Biophys. Acta* 633: 400-409. These patients typically have a large accumulation of uracil and thymine in their cerebrospinal fluid (*see*, Bakkeren *et al.*, *supra.*) and urine (*see*, Berger *et al.*, *supra.*; Bakkeren *et al.*, *supra.*; Brockstedt *et al.*, *supra.*; and Fleming *et al.* (1992) *Cancer Res.* 52: 2899-2902).

Familial studies suggest that DPD deficiency follows an autosomal recessive pattern of inheritance. *See*, Diasio *et al.*, (1988) *supra*. Up to three percent of the general human population are estimated to be putative heterozygotes for DPD deficiency, as determined by enzymatic activity in lymphocytes. *See*, Milano and Eteinne (1994) *Pharmacogenetics*. This suggests that the frequency of homozygotes for DPD deficiency may be as high as one person per thousand.

DPD has been purified from liver tissue of rats (*see*, Shiotani and Weber (1981) *J. Biol. Chem.* 256: 219-224; Fujimoto *et al.* (1991); *J. Nutr. Sci. Vitaminol.* 37:

89-98], pig [Podschun *et al.* (1989) *Eur. J. Biochem.* 185: 219-224), cattle (*see*, Porter *et al.* (1991) *J. Biol. Chem.* 266: 19988-19994), and humans (*see*, Lu *et al.* (1992) *J. Biol. Chem.* 267: 1702-1709). The pig enzyme contains flavins and iron-sulfur prosthetic groups and exists as a homodimer with a monomer Mr of about 107,000 (*see*, Podschun *et al.*, *supra.*). Since the enzyme exhibits a nonclassical two-site ping-pong mechanism, it appears to have distinct binding sites for NADPH/NADP and uracil/5,6-dihydrouracil (*see*, Podschun *et al.* (1990) *J. Biol. Chem.* 265: 12966-12972). An acid-base catalytic mechanism has been proposed for DPD (*see*, Podschun *et al.* (1993) *J. Biol. Chem.* 268: 3407-3413).

The DPD cDNA is described in copending U.S. application SN 08/304,309. Recently, DPD mRNA from patients lacking dihydrompyrimidine activity was found to lack an exon which encodes a 165 base pair sequence found in the wild-type DPD cDNA. *See*, Meinsma *et al.* (1995) *DNA and Cell Biology* 14(1): 1-6.

Because an undetected DPD deficiency poses a significant danger to a cancer patient who is being treated with 5-FU, a great need exists for a simple and accurate test for DPD deficiency. Such a test will also facilitate diagnosis of disorders that are associated with DPD deficiency, such as uraciluria. The present invention provides such a test, thus fulfilling these and other needs.

SUMMARY OF THE INVENTION

Particular mutations in the dihydropyrimidine dehydrogenase gene are described herein which lead to loss of dihydropyrimidine dehydrogenase activity. The mutations result in a loss of the amino acids from 581-635 of the protein encoded by the gene due to a splicing defect. The splicing defect results in the loss of an exon encoding the missing amino acids. The assays and compositions of the invention detect the splicing defect in the genomic DNA which results in the loss of the exon.

A variety of assays for detecting splicing defect mutations in patients are provided. The assays determine whether RNA encoded by genomic DNA is competent to be spliced to produce mRNA with nucleic acids encoding the exon which corresponds to amino acids from 581-635 of the wild-type protein. For example, sequencing, PCR, LCR, and oligonucleotide array based assays are used to detect the mutations.

In ne class of embodiments, the methods comprise the step of determining whether a nucleic acid encoding an mRNA for the dihydropyrimidine dehydrogenase gene has an exon corresponding to amino acids 581-635 of a corresponding wild-type dihydropyrimidine dehydrogenase mRNA. This determination is performed in sequencing, PCR, LCR, and oligonucleotide array based hybridization assays. For example, in one class of PCR based assays, the intronic genomic DNA encoding the dihydropyrimidine dehydrogenase in the region flanking the exon which encodes amino acids 581-635 is hybridized to a set of PCR primers for amplification and analysis of the intron-exon splice boundaries.

Example primers which are used for amplifying the splice regions include DELF1 and DELR1. Similar primers which hybridize to the same sites, or to sites proximal to the primer binding sites are also used.

One particular mutation which results in an abnormal dihydropyrimidine dehydrogenase gene phenotype is the conversion of a G to an A residue at the 3' GT splice site of the exon which encodes amino acids 581-635 of the corresponding wild-type protein. The 3' wild-type splice site is recognized by restriction endonucleases which recognize the Mae II cleavage site. The ability of the cleavage site to be cleaved by restriction endonucleases which recognize the Mae II site is used to distinguish wild-type from abnormal dihydropyrimidine dehydrogenase genomic DNAs. For example, a region including the mutation is amplified by PCR, and the PCR products cleaved by Mae II. This results in the cleavage of nucleic acids amplified by the wild-type gene, but mutant sequences are not amplified.

In addition to PCR detection methods, nucleic acid arrays are used to discriminate single base-pair mismatches, or to directly sequence DNA by hybridization to arrays. The sequence of the splice site is also determined by standard or PCR sequencing of the site, *e.g.*, using primers which flank the site in a pol I or *taq* PCR based sequencing assay.

In addition to methods and compositions for abnormal dihydropyrimidine dehydrogenase gene detection, the invention provides kits for practicing the methods. Typically, the kit contains a first PCR primer which binds to DNA 3' of a splice site in the genomic DNA for dihydropyrimidine dehydrogenase gene for an exon encoding amino acids 581-635, and a second PCR primer which binds to DNA 5' of a splice site in the

genomic DNA f r dihydropyrimidine dehydrogenase gene for an exon encoding amino acids 581-635. The kit optionally contains other components such as instructions for the detection of an abnormal dihydropyrimidine dehydrogenase gene, restriction enzymes such as Mae II, PCR or other *in vitro* amplification reagents (buffers, enzymes, salts and the like), oligonucleotide array detection chips and the like.

DETAILED DESCRIPTION OF THE DRAWING

Figure 1 provides genomic sequence from the DPD gene in the region of a splicing mutant which leads to the loss of an exon in the mRNA which encodes amino acids 581-635. The primer binding sites for DELF-1 and DELR1 are indicated. The primer binding site for DPD15F and DPD15R (*see*, Meinsma *et al.* (1995) *DNA and Cell Biology* 14(1): 1-6) are also indicated. The Mae II site at the 3' splice junction is indicated. In a mutant form, the G nucleotide at the 3' splice junction (residue 434) is mutated to an A nucleotide.

Figure 2 provides further details of PCR reaction components for amplifying the region of a splicing site mutant. The exon is underlined. The splice site which is polymorphic is in bold text.

Figure 3 is a schematic of the splicing product of the mutant DPD gene.

Figure 4 is a schematic representation of an example strategy for amplifying an exon deleted in mutant Dpd mRNA from the genomic DNA for the gene. The restriction endonuclease Mae II produces fragments of 278 bp and 131 bp in the wild-type allele. A point mutation at the donor splicing consensus sequence (GT to AT) destroys the restriction site, and thus, the mutant form remains undigested (409 bp).

DEFINITIONS

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Singleton *et al.* (1994) Dictionary of Microbiology and Molecular Biology, second edition, John Wiley and Sons (New York); Walker (ed) (1988) The Cambridge Dictionary of Science and Technology, The press syndicate of the University of Cambridge, NY; and Hale and Marham (1991) The Harper Collins Dictionary of Biology Harper Perennial, NY provide one of skill with a general

dictionary of many of the terms used in this invention. Although any methods and materials similar r equivalent to those described herein can be used in the practice or testing of the present invention, preferred methods and materials are described. For purposes of the present invention, the following terms are defined below.

The term "nucleic acid" refers to deoxyribonucleotides, ribonucleotides, and polymers thereof in either single- or double-stranded form, and unless specifically limited, encompasses known analogues of natural nucleotides that hybridize to nucleic acids in a manner similar to naturally occurring nucleotides. Unless otherwise indicated, a particular reference nucleic acid sequence implicitly encompasses conservatively modified variants thereof (*i.e.*, sequence variants that hybridize with the reference nucleic acid, that encode the same amino acid sequence as the reference nucleic acid, or that encode amino acid sequences containing conservative amino acid substitutions as compared to the amino acid sequence encoded by the reference nucleic acid), and complementary sequences and as well as the sequence explicitly indicated.

The term "wild type," when applied to a gene, nucleic acid or gene product, especially a protein and/or biological property, denotes a gene, gene product, protein, or biological property predominantly found in nature. "Mutant" proteins, nucleic acids and the like occur less frequently and are often associated with loss or alteration of biological activity and a diseased state.

The term "dihydropyrimidine dehydrogenase gene" refers to a naturally occurring nucleic acid that encodes a protein that has dihydropyrimidine dehydrogenase activity, or that is specifically recognized by antibodies specific for dihydropyrimidine dehydrogenase. The term encompasses both wild-type and mutant genes.

"Intron-exon boundary" denotes a sequence of a gene present in or derived from a genomic DNA, wherein the sequence comprises intron sequences that are absent in a mature mRNA transcript, contiguous to exon regions that are present in the mature mRNA transcript.

A "splicing defect" is a failure to correctly remove an intron or to correctly connect two exons in the mature mRNA transcript, which often are the result of sequence mutations in the intron-exon boundary.

"Stringency" relates to the conditions under which hybridization and annealing take place. Low stringency hybridization and annealing conditions permit the

annealing of complementary nucleic acids that contain mismatched nucleic acids. As the stringency is raised, annealing f sequences containing mismatched nucleic acids is disfavored. Conditions which result in low or high stringency levels are known in the art (*e.g.*, increasing the annealing temperature raises the stringency).

An extensive guide to the hybridization of nucleic acids is found in Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Acid Probes* part I chapter 2 "overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York. Generally, highly stringent hybridization and wash conditions are selected to be about 5° C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and ph. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Very stringent conditions are selected to be equal to the T_m for a particular probe.

Methods of alignment of sequences for comparison and determination of whether two sequences bind the same target are well known in the art. Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman (1981) *Adv. Appl. Math.* 2: 482; by the homology alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48: 443; by the search for similarity method of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444; by computerized implementations of these algorithms (including, but not limited to CLUSTAL in the PC/Gene program by Intelligenetics, Mountain View, California, GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, Wisconsin, USA); the CLUSTAL program is well described by Higgins and Sharp (1988) *Gene*, 73: 237-244 and Higgins and Sharp (1989) *CABIOS* 5: 151-153; Corpet, *et al.* (1988) *Nucleic Acids Research* 16, 10881-90; Huang, *et al.* (1992) *Computer Applications in the Biosciences* 8, 155-65, and Pearson, *et al.* (1994) *Methods in Molecular Biology* 24, 307-31. Alignment is also often performed by inspection and manual alignment.

A "restriction endonuclease which recognizes a Mae II cleavage site" is a DNA endonuclease that recognizes and cleaves a nucleic acid which is cleaved by the restriction endonuclease Mae II (*see also*, Figure 1).

A "PCR primer" is one of a pair of oligonucleotides complementary to a subregion of a nucleic acid that, when annealed to the nucleic acid, generates a double stranded sequence which supports nucleic acid synthesis by primer extension in a PCR reaction.

DETAILED DISCUSSION OF THE INVENTION

The present invention provides compositions, methods and kits for detecting point mutations in DPD nucleic acid samples. Genetic amplification techniques are used to detect point mutations in a splicing region of the DPD gene which results in the deletion of an exon in the DPD gene product, and causes one form of DPDD (DPD deficiency).

Isolation of nucleic acid samples

To amplify a target intron-exon boundary in a sample, the nucleic acid encoding the sequence is made accessible to the components of the amplification system. In general, this accessibility is ensured by isolating the nucleic acids from the sample, however, isolation is optional (methods for amplifying nucleic acids, *e.g.*, by PCR from whole cells are known and appropriate). A variety of techniques for extracting nucleic acids from biological samples are known in the art. For example, see those described in Rotbart *et al.*, 1989, in *PCR Technology* (Erlich ed., Stockton Press, New York) and Han *et al.*, 1987, *Biochemistry* 26:1617-1625. The methods described by Fries *et al.*, *Am. J. Med. Genet.*, 46:363-368 (1993), are also useful.

Nucleic acids are isolated from cell or tissue samples from patients, and from cell culture. The culture of cells used in conjunction with the present invention, including cell lines and cultured cells from tissue or blood samples, including stem cells is well known in the art. Freshney (*Culture of Animal Cells, a Manual of Basic Technique, third edition* Wiley-Liss, New York (1994)) and the references cited therein provides a general guide to the culture of cells. *See also*, Kuchler *et al.* (1977) *Biochemical Methods in Cell Culture and Virology*, Kuchler, R.J., Dowden, Hutchinson and Ross, Inc, and Inaba *et al.* (1992) *J. Exp. Med.* 176, 1693-1702.

The methods of the present invention are employed for analyzing nucleic acid from a variety of different tissues, such as tissues samples (*e.g.*, fetal and amniotic fluid tissue) and patient-derived tissue culture cell lines (*e.g.*, fibroblast cell lines). Where

blood cells are used, whole blood should be drawn with an anticoagulant in a sealed vacuum tube kept separated from other samples and handled with clean gl ves. For best results, blood is processed immediately after collection; if this is impractical, it is kept in a sealed container until use. Cells in other physiological fluids, such as amniotic fluid, are also optionally assayed. When using any of these fluids, the cells in the fluid are typically separated from the fluid component by centrifugation.

Several methods are particularly well suited for use in amplifying target regions. The choice between these methods is typically governed by the size of the sample which is available for testing. One method is crude extraction which is useful for relatively large samples. Briefly, mononuclear cells from samples of blood, buffy coat, amniocytes from amniotic fluid, cultured chorionic villus cells, or the like are isolated by standard procedures (*see*, Freshney, *supra*). If testing is not performed immediately after sample collection, aliquots of cells can be pelletted in sterile Eppendorf tubes and the dry pellet frozen at -20°C or colder until use.

The blood cells are resuspended (typically from 10 to about 10^6 nucleated cells per 100 μl to 1 ml) in a buffer of 50 mM Tris-HCl (pH 8.3), 50 mM KCl 1.5 mM $MgCl_2$, 0.5% Tween 20, 0.5% NP40 supplemented with 100 μg/ml of proteinase K. After incubating at 56°C for 2 hr, the cells are heated to 95°C for 10 min to inactivate the proteinase K and immediately moved to wet ice (snap-cool). If gross aggregates are present, another cycle of digestion in the same buffer should be undertaken. Ten μl of this extract is generally used for amplification.

When extracting DNA from tissues, *e.g.*, chorionic villus cells or confluent cultured cells, the amount of the above mentioned buffer with proteinase K varies according to the size of the tissue sample. The extract is incubated for 4-10 hrs at 50°-60°C and then at 95°C for 10 minutes to inactivate the proteinase. During longer incubations, fresh proteinase K should be added after about 4 hr at the original concentration.

When the sample contains a small number of cells, extraction may be accomplished, *e.g.*, by methods as described in Higuchi, "Simple and Rapid Preparation of Samples for PCR", in *PCR Technology*, Ehrlich, H.A. (ed.), Stockton Press, New York, which is incorporated herein by reference.

A relatively easy procedure for extracting DNA for amplificati n is a "salting out" procedure adapted from the method described by Miller *et al.*, *Nucleic Acids Res.*, 16:1215 (1988). Mononuclear cells are separated on a Ficoll-Hypaque gradient. The cells are resuspended in 3 ml of lysis buffer (10 mM Tris-HCl, 400 mM NaCl, 2 mM Na_2EDTA, pH 8.2). Fifty μl of a 20 mg/ml solution of proteinase K and 150 μl of a 20% SDS solution are added to the cells and then incubated at 37°C overnight. Rocking the tubes during incubation will improve the digestion of the sample. If the proteinase K digestion is incomplete after overnight incubation (fragments are still visible), an additional 50 μl of the 20 mg/ml proteinase K solution is mixed in the solution and incubated for another night at 37°C on a gently rocking or rotating platform. Following adequate digestion, one ml of a 6M NaCl solution is added to the sample and vigorously mixed. The resulting solution is centrifuged for 15 minutes at 3000 rpm. The pellet contains the precipitated cellular proteins, while the supernatant contains the DNA. The supernatant is removed to a 15 ml tube that contains 4 ml of isopropanol. The contents of the tube are mixed gently until the water and the alcohol phases have mixed and a white DNA precipitate has formed. The DNA precipitate is removed and dipped in a solution of 70% ethanol and mixed gently. The DNA precipitate is removed from the ethanol and air dried. The precipitate is placed in distilled water and dissolved.

Kits are also commercially available for the extraction of high-molecular weight (*i.e.*, genomic) DNA. These kits include Genomic Isolation Kit A.S.A.P. (Boehringer Mannheim, Indianapolis, IN), Genomic DNA Isolation System (GIBCO BRL, Gaithersburg, MD), Elu-Quik DNA Purification Kit (Schleicher & Schuell, Keene, NH), DNA Extraction Kit (Stratagene, La Jolla, CA), TurboGen Isolation Kit (Invitrogen, San Diego, CA), and the like. Use of these kits according to the manufacturer's instructions is generally acceptable for purification of DNA when practicing the methods of the present invention.

The concentration and purity of the extracted DNA may be determined by spectrophotometric analysis of the absorbance of a diluted aliquot at 260 nm and 280 nm. Either 10 μl of crude extract, or 1 μg of purified DNA by the alternate methods are used for amplification.

PCR Amplification of the Intron-Exon Boundaries f the DPD gene

The nucleic acids at intron-exon boundaries are typically amplified when determining whether a splicing defect is present in a sample. In a preferred embodiment, amplification is performed by the PCR method. Although the PCR process is well known in the art (*see*, U.S. Patent Nos. 4,683,195; 4,683,202; and 4,965,188, each of which is incorporated herein by reference) and although commercial vendors, such as Roche Molecular Systems, sell PCR reagents and publish PCR protocols, some general PCR information is provided below for purposes of clarity and full understanding of the invention for those unfamiliar with the PCR process.

In brief, a DNA sample is heat denatured to separate the individual complementary strands and annealed in the presence a pair of oligonucleotide primers which are complementary to sequences present on opposite stands and located on either side of the sequence to be detected. The PCR mixture is then subjected to repeated cycles of DNA synthesis, denaturation and reannealing.

The first step of each cycle of the PCR involves the separation of the nucleic acid duplex formed by the primer extension. Once the strands are separated, the next step in PCR involves hybridizing the separated strands with primers that flank the target sequence. The primers are then extended to form complementary copies of the target strands. For successful PCR amplification, the primers are designed so that the position at which each primer hybridizes along a duplex sequence is such that an extension product synthesized from one primer, when separated from the template (complement), serves as a template for the extension of the other primer. The cycle of denaturation, hybridization, and extension is repeated as many times as necessary to obtain the desired amount of amplified nucleic acid.

In a preferred embodiment of the PCR process, strand separation is achieved by heating the reaction to a sufficiently high temperature for an sufficient time to cause the denaturation of the duplex but not to cause an irreversible denaturation of the polymerase (see U.S. Patent No. 4,965,188). Typical heat denaturation involves temperatures ranging from about 80°C to 105°C for times ranging from seconds to minutes. Strand separation, however, can be accomplished by any suitable denaturing method including physical, chemical, or enzymatic means. Strand separation may be induced by a helicase, for example, or an enzyme capable of exhibiting helicase activity. For example, the enzyme

RecA has helicase activity in the presence of ATP. The reaction conditions suitable f r strand separation by helicases are known in the art (see Kuhn Hoffman-Berling, 1978, *CSH-Quantitative Biology* 43:63-67; and Radding, 1982, *Ann. Rev. Genetics* 16:405-436, both of ·vhich are incorporated herein by reference).

Template-dependent extension of primers in PCR is catalyzed by a polymerizing agent in the presence of adequate amounts of four deoxyribonucleoside triphosphates (typically dATP, dGTP, dCTP, and dTTP) in a reaction medium comprised of the appropriate salts, metal cations, and pH buffering system. Suitable polymerizing agents are enzymes known to catalyze template-dependent DNA synthesis. In some instances, DPD-encoding RNA may be used as the initial template for primer extension is RNA. Polymerizing agents suitable for synthesizing a complementary, DNA (cDNA) sequence from the RNA template are reverse transcriptase (RT), such as avian myeloblastosis virus RT, Moloney murine leukemia virus RT, or *Thermus thermophilus* (*Tth*) DNA polymerase, a thermostable DNA polymerase with reverse transcriptase activity marketed by Roche Molecular Systems Typically, the RNA template is heat degraded during the first denaturation step after the initial reverse transcription step leaving only DNA template. Suitable polymerases for use with a DNA template include, for example, *E. coli* DNA polymerase I or its Klenow fragment, T_4 DNA polymerase, *Tth* polymerase, and *Taq* polymerase, a heat-stable DNA polymerase isolated from *Thermus aquaticus* and commercially available from Roche Molecular Systems The latter enzyme is widely used in the amplification and sequencing of nucleic acids. The reaction conditions for using *Taq* polymerase are known in the art and are described in Innis, *supra*.

When RNA is amplified, an initial reverse transcription (RT) step is carried out to create a DNA copy (cDNA) of the RNA. PCT patent publication No. WO 91/09944, published July 11, 1991, incorporated herein by reference, describes high-temperature reverse transcription by a thermostable polymerase that also functions in PCR amplification. High-temperature RT provides greater primer specificity and improved efficiency. A "homogeneous RT-PCR" in which the same primers and polymerase suffice for both the reverse transcription and the PCR amplification steps, and the reaction conditions are optimized so that both reactions occur without a change of reagents is also available. *Thermus thermophilus* DNA polymerase, a thermostable DNA polymerase that can functi n as a reverse transcriptase, is used for all primer extensi n

steps, regardless of template. Both processes can be done without having to open the tube to change or add reagents; only the temperature profile is adjusted between the first cycle (RNA template) and the rest of the amplification cycles (DNA template).

The PCR method can be performed in a step-wise fashion, where after each step new reagents are added, or in a fashion where all of the reagents are added simultaneously, or in a partial step-wise fashion, where fresh or different reagents are added after a given number of steps. For example, if strand separation is induced by heat, and the polymerase is heat-sensitive, then the polymerase is added after every round of strand separation. However, if, for example, a helicase is used for denaturation, or if a thermostable polymerase is used for extension, then all of the reagents are added initially, or, alternatively, if molar ratios of reagents are of consequence to the reaction, the reagents may be replenished periodically as they are depleted by the synthetic reaction.

Those skilled in the art will know that the PCR process is most usually carried out as an automated process with a thermostable enzyme. In this process, the temperature of the reaction mixture is cycled through a denaturing region, a primer annealing region, and an extension reaction region. Alternatively, the annealing and extension temperature can be the same. Reverse transcriptase-PCR uses such a two-step temperature cycling. A machine specifically adapted for use with a thermostable enzyme is commercially available from Roche Molecular Systems.

Those skilled in the art will also be aware of the problems of contamination of a PCR by the amplified nucleic acid from previous reactions and nonspecific amplification. Methods to reduce these problems are provided in PCT patent application Serial No. 91/05210, filed July 23, 1991, incorporated herein by reference. The method allows the enzymatic degradation of any amplified DNA from previous reactions and reduces nonspecific amplification. The PCR amplification is carried out in the presence of dUTP instead of dTTP. The resulting double-stranded, uracil-containing product is subject to degradation by uracil N-glycosylase (UNG), whereas normal thymine-containing DNA is not degraded by UNG. Adding UNG to the amplification reaction mixture before the amplification is started degrades all uracil-containing DNA that might serve as target. Because the only source of uracil-containing DNA is the amplified product of a previous reaction, this method effectively sterilizes the reaction mixture, eliminating the problem of contamination from previous reactions (carry-over). UNG itself is rendered temporarily

inactive by heat, so the denaturation steps in the amplification procedure also serve to inactivate the UNG. New amplification products, therefore, though incorporating uracil, are formed in an effectively UNG-free environment and are not degraded.

Those practicing the present invention should note that, although the preferred embodiment incorporates PCR amplification, amplification of target sequences in a sample may be accomplished by any known method, such as ligase chain reaction (LCR), transcription amplification, and self-sustained sequence replication, each of which provides sufficient amplification so that the target sequence can be detected by nucleic acid hybridization to a probe. Persons of skill will appreciate that in methods such as LCR, primers that are complementary to the specific polymorphism or mutation are used. In this instance amplification occurs when the polymorphism (*i.e.*, point mutation) is present in the nucleic acid sample.

Alternatively, methods that amplify the probe to detectable levels can be used, such as Qβ-replicase amplification. The term "probe" encompasses, *inter alia*, the sequence specific oligonucleotides used in the above procedures; for instance, the two or more oligonucleotides used in LCR are "probes" for purposes of the present invention, even though some embodiments of LCR only require ligation of the probes to indicate the presence of an allele.

Examples of techniques sufficient to direct persons of skill through such *in vitro* amplification methods, including the polymerase chain reaction (PCR) the ligase chain reaction (LCR), Qβ-replicase amplification and other RNA polymerase mediated techniques (*e.g.*, NASBA) are found in Berger and Kimmel, *Guide to Molecular Cloning Techniques*, Methods in Enzymology volume 152 Academic Press, Inc., San Diego, CA (Berger); Sambrook et al. (1989) Molecular Cloning - A Laboratory Manual (2nd ed.) Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor Press, NY, (Sambrook); and Current Protocols in Molecular Biology, F.M. Ausubel et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (1994 Supplement) (Ausubel), and in Mullis *et al.*, (1987) U.S. Patent No. 4,683,202; *PCR Protocols A Guide to Methods and Applications* (Innis *et al.* eds) Academic Press Inc. San Diego, CA (1990) (Innis); Arnheim & Levinson (October 1, 1990) *C&EN* 36-47; *The Journal Of NIH Research* (1991) 3, 81-94; (Kwoh *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86, 1173; Guatelli *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87, 1874; Lomell *et al.*

(1989) *J. Clin. Chem* 35, 1826; Landegren *et al.*, (1988) *Science* 241, 1077-1080; Van Brunt (1990) *Biotechnology* 8, 291-294; Wu and Wallace, (1989) *Gene* 4, 560; Barringer *et al.* (1990) *Gene* 89, 117, and Sooknanan and Malek (1995) *Biotechnology* 13: 563-564. Improved methods of cloning *in vitro* amplified nucleic acids are described in Wallace *et al.*, U.S. Pat. No. 5,426,039.

The present invention provides, *inter alia*, a polymorphic form of DPD. In particular, the invention provides a DPDD polymorphism wherein the G at position 434 of the gene sequence shown in Figure 1 is replaced by an A.

This particular polymorphism is detected by a variety of amplification techniques, preferably PCR as described *supra*. To detect this polymorphism by PCR, the PCR reaction is performed in the presence of primers that are complimentary to opposite strands of the genomic DNA, wherein the complementary sequences are located on either side of the point mutation. The precise sequences recognized by the primers are not critical. Typically, any pair of primers can be used as long as they (1) bracket the polymorphism, (2) are reasonably near to the polymorphism (while the primer binding sequence may be as far from the polymorphism as can support a PCR reaction, *i.e.*, 1 to about 10 kb, it is preferable that the binding sequence be within about 500 nucleotides or less, and more preferable that the binding sequence be within 100 nucleotides of the exon-intron boundary to be assayed), and (3) bind the primers with an adequate degree of specificity. It is preferable that the sequence be unique to the gene of interest. Such sequences are identified by comparing sequences as described herein. Smaller primers have a higher probability of recognizing sites outside of the desired binding site, whereas very large primers are more expensive to make; generally, a primer of about 15-20 nucleotides is adequate, and therefore preferred.

Example primers used herein are DELF1, which hybridizes to nucleic acids corresponding to nucleotide 154 to nucleotide 175 of the DPD gene sequence, and DELR1, which hybridizes to nucleic acids corresponding to nucleotide 563 to nucleotide 542 of the DPD gene sequence as described in Figure 1. It is important to note that other 15-20 nucleotide sequences within 500 nucleotides of the point mutation are also acceptable. Given the sequence of the affected exon, adjacent introns, and adjacent additional exons provided herein, selecting and making alternate primers is within the skill of a skilled practitioner in the relevant art.

The present invention also provides kits for the detection of genetic polymorphisms or mutations associated with DPDD. The kits comprise a vial containing amplification primers that span a DPDD-associated polymorphism or mutation in the gene encoding DPD. For example, the vial optionally contains DELF-1 and DELR-1. The kits optionally contain a vial containing a thermostable polymerase, genetic size markers for gels, amplification reagents, instructions and the like.

Sequence analysis of Amplification Products

A variety of methods are employed to analyze the nucleotide sequence of the amplification products. Several techniques for detecting point mutations following amplification by PCR have been described in Chehab *et al.*, *Methods in Enzymology*, 216:135-143 (1992); Maggio *et al.*, *Blood*, 81(1):239-242 (1993); Cai and Kan, *Journal of Clinical Investigation*, 85(2):550-553 (1990) and Cai *et al.*, *Blood*, 73:372-374 (1989).

One particularly useful technique is analysis of restriction enzyme sites following amplification. In this method, amplified nucleic acid segments are subjected to digestion by restriction enzymes. Identification of differences in restriction enzyme digestion between corresponding amplified segments in different individuals identifies a point mutation. Differences in the restriction enzyme digestion is commonly determined by measuring the size of restriction fragments by electrophoresis and observing differences in the electrophoretic patterns. Generally, the sizes of the restriction fragments is determined by standard gel electrophoresis techniques as described in Sambrook, and, *e.g.*, in Polymeropoulos *et al.*, *Genomics*, 12:492-496 (1992).

The size of the amplified segments obtained from affected and normal individuals and digested with appropriate restriction enzymes are analyzed on agarose or polyacrylamide gels. Because of the high discrimination of the polyacrylamide gel electrophoresis, differences of small magnitude are easily detected. Other mutations resulting in DPDD-related polymorphisms of DPD encoding genes also add unique restriction sites to the gene that are determined by sequencing DPDD-related nucleic acid sequences and comparing them to normal sequences.

Another useful method of identifying point mutations in PCR amplification products employs oligonucleotide probes specific for different sequences. The ligonucleotide probes are mixed with amplification products under hybridization

conditions. Probes are either RNA or DNA oligonucleotides and optionally contain not nly naturally occurring nucleotides but also analogs such as dig xygenin dCTP, biotin dCTP, 7-azaguanosine, azidothymidine, inosine, or uridine. The advantage of using nucleic acids comprising analogs include selective stability, resistance to nuclease activity, ease of signal attachment, increased protection from extraneous contamination and an increased number of probe-specific colored labels. For instance, in preferred embodiments, oligonucleotide arrays are used for the detection of specific point mutations as described below.

Probes are typically derived from cloned nucleic acids, or are synthesized chemically. When cloned, the isolated nucleic acid fragments are typically inserted into a replication vector, such as lambda phage, pBR322, M13, pJB8, c2RB, pcos1EMBL, or vectors containing the SP6 or T7 promoter and cloned as a library in a bacterial host. General probe cloning procedures are described in Arrand J.E., *Nucleic Acid Hybridization A Practical Approach*, Hames B.D., Higgins, S.J., Eds., IRL Press 1985, pp. 17-45 and Sambrook, J., Fritsch, E.F., Maniatis, T., *Molecular Cloning A Laboratory Manual*, Cold Spring Harbor Press, 1989, pp. 2.1-3.58, both of which are incorporated herein by reference.

Oligonucleotide probes and primers are synthesized chemically with or without fluorochromes, chemically active groups on nucleotides, or labeling enzymes using commercially available methods and devices like the Model 380B DNA synthesizer from Applied Biosystems, Foster City, California, using reagents supplied by the same company. Oligonucleotides for use as probes, *e.g.*, in *in vitro* amplification methods, or for use as gene probes are typically synthesized chemically according to the solid phase phosphoramidite triester method described by Beaucage and Caruthers (1981), *Tetrahedron Letts.*, 22(20):1859-1862, *e.g.*, using an automated synthesizer, as described in Needham-VanDevanter *et al.* (1984) *Nucleic Acids Res.*, 12:6159-6168. Oligonucleotides can also be custom made and ordered from a variety of commercial sources known to persons of skill. Purification of oligonucleotides, where necessary, is typically performed by either native acrylamide gel electrophoresis or by anion-exchange HPLC as described in Pearson and Regnier (1983) *J. Chrom.* 255:137-149. The sequence of the synthetic oligonucleotides can be verified using the chemical degradation method of Maxam and

Gilbert (1980) in Grossman and Moldave (eds.) Academic Press, New York, *Methods in Enzymology* 65:499-560.

Oligonucleotide probes and primers are selected using commercially available computer programs to compare known DNA sequences from gene sequences found in gene libraries, such as Genebank and EMBL, and the sequences described herein. The programs identify unique nucleotide sequences within the gene of interest. One such program is Eugene. Oligonucleotide sequences for PCR of a unique genomic DNA such as a chromosome subsequence are chosen optimally by choosing sequences according to previously established protocols or by computer programs that choose the degree of homology desired along with the length of the probe. Sequences are chosen to avoid technical problems such as primer dimers resulting from amplification of hybridized primers.

Primers and probes are optionally labeled with fluorophores or enzymes that generate colored products. This allows simultaneous use of probes to different DPDD-related polymorphisms or mutations. Identification of hybridization of a specifically labelled primer provides a means for determining which polymorphism or mutation is present in the nucleic acid of the sample. The primers used in the assay are labeled with more than one distinguishable fluorescent or pigment color. Primers are labeled with Texas red, rhodamine and its derivatives, fluorescein and its derivatives, dansyl, umbelliferone and the like or with horse radish peroxidase, alkaline phosphatase, biotin, avidin, or the like.

Primers and probes are labeled directly or indirectly. The common indirect labeling schemes covalently bind a ligand to the nucleotide and prepare labeled probe by incorporating the ligand using random priming or nick translation. The ligand then binds an anti-ligand which is covalently bound to a label. Ligands and anti-ligands vary widely. When a ligand has an anti-ligand, *e.g.*, biotin, thyroxine, or cortisol, the ligand is used in conjunction with the labelled naturally-occurring anti-ligand. Alternatively, a hapten or antigen may be used in combination with an antibody, which is optionally labeled.

Sequence specific oligonucleotide probes hybridize specifically with a particular segment of the target polymorphism or mutation amplification products and have destabilizing mismatches with the sequences from other polymorphisms or mutations. Under sufficiently stringent hybridization conditions, the probes hybridize specifically only

to exactly compl mentary sequences. The stringency of the hybridization conditi ns can be relaxed to tolerate varying am unts of sequence mismatch. Detection of the amplified product utilizes this sequence-specific hybridization to insure detection of only the correct amplified target, thereby decreasing the chance of a false positive caused by the presence of homologous sequences from related polymorphisms or mutations.

Specific DPD-encoding polymorphisms or mutations are also identified by sequencing the amplification products or restriction fragments thereof. Sequencing is performed by a variety of methods well known in the art. For example, the sequence of the amplified nucleic acid segments may be determined by the Maxam-Gilbert chemical degradation method as described in Sambrook. Generally, Sanger dideoxy-mediated sequencing is employed as described in Sambrook, or sequencing by hybridization is performed as described below.

Detection of Splicing Site Defects Using Nucleic acid Arrays

In one preferred class of embodiments, splicing site defects at intron-exon boundaries are detected by hybridization of amplification products which include the splicing site to oligonucleotide arrays which discriminate single base-pair mismatches. In this embodiment, primers are used to amplify an intron-exon boundary in a PCR reaction, resulting in PCR amplicons which comprise the intron-exon boundary. The sequence of the entire PCR amplicon, or any subsequence thereof can be determined by labeling the PCR amplicon (typically with biotin or a fluorescent label) and hybridization to an array of oligonucleotide probes. In these hybridization methods single base pair mismatches in labeled nucleic acids to probes in the array are distinguished.

Preferably in this class of embodiments, the oligonucleotide arrays are designed to sequence nucleic acids at the intron-exon boundary. More preferably, the arrays are designed to discriminate whether a particular nucleotide is altered relative to the wild-type sequence. This is done by constructing an array with two or more oligonucleotide probe sets which differ by a single nucleotide. Hybridization to the known probe sequence by a target nucleic acid under conditions where a single mismatch does not bind indicates the presence of a fully complementary nucleic acid.

Sequencing by hybridization to arrays of oligonucleotides is described in U.S. Patent N . 5,202,231, to Drmanac *et al.* and, *e.g.*, in Drmanac *et al.* (1989)

Genomics 4:114-128. Methods of constructing and designing arrays for sequencing and detection of single nucleotide alterations is known in the art. The development of very large scale immobilized polymer synthesis (VLSIPS™) technology provides methods for arranging large numbers of oligonucleotide probes for the detection and sequencing of nucleic acids in very small arrays. *See*, WO 90/15070 and 92/10092; Pirrung *et al.*, U.S. Patent No. 5,143,854 (see also PCT Application No. WO 90/15070); McGall *et al.*, U.S. Patent No. 5,412,087; and US Pat. No. 5,384,261. *See also*, Fodor *et al.* (1991) Science, 251: 767-777 and Sheldon *et al.* (1993) Clinical Chemistry 39(4): 718-719. The oligonucleotide arrays are typically placed on a solid surface such as a glass slide with an area less than 1 inch^2, although much larger surfaces are optionally used.

Mechanical and Light directed oligonucleotide array construction methods are used for the construction of oligonucleotide arrays. Light directed methods are the most common, and are found, *e.g.*, in U.S. Patent No. 5,143,854. The light directed methods discussed in the '854 patent typically proceed by activating predefined regions of a substrate or solid support and then contacting the substrate with a preselected monomer solution. The predefined regions are activated with a light source, typically shown through a photolithographic mask. Other regions of the substrate remain inactive because they are blocked by the mask from illumination. Thus, a light pattern defines which regions of the substrate react with a given nucleic acid reagent. By repeatedly activating different sets of predefined regions and contacting different reagent solutions with the substrate, a diverse array of oligonucleotides is produced on the substrate. Other steps, such as washing unreacted reagent solutions from the substrate, are used as necessary.

The surface of a solid support is typically modified with linking groups having photolabile protecting groups and illuminated through a photolithographic mask, yielding reactive hydroxyl groups in the illuminated regions. For instance, during oligonucleotide synthesis, a 3'-O- phosphoramidite (or other nucleic acid synthesis reagent) activated deoxynucleoside (protected at the 5'-hydroxyl with a photolabile group) is then presented to the surface and coupling occurs at sites that were exposed to light in the previous step. Following capping, and oxidation, the substrate is rinsed and the surface illuminated through a second mask, to expose additional hydroxyl groups for coupling. A second 5'-protected, 3'-O-phosphoramidite activated deoxynucleoside (or other monomer as appropriate) is then presented to the resulting array. The selective photodeprotection

and coupling cycles are repeated until the desired set of ligonucleotides (or other polymers) is produced.

The PCR amplicons detected on the arrays are labeled with a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, or chemical means. For example, useful labels include 32P, 35S, fluorescent dyes, chromophores, electron-dense reagents, enzymes (e.g., as commonly used in an ELISA), biotin, dioxigenin, or haptens and proteins for which antisera or monoclonal antibodies are available. In preferred embodiments, the label is detectable spectroscopically, *i.e.*, is chromogenic. Suitable chromogens include molecules and compounds which absorb light in a distinctive range of wavelengths so that a color may be observed, or emit light when irradiated with radiation of a particular wavelength or wavelength range (*e.g.*, a fluorescent label).

EXAMPLES

The following examples are offered by way of illustration and not limitation. One of skill will readily recognize a variety of parameters and conditions which can be changed or modified to yield essentially identical results.

The following example pinpoints a polymorphism or "mutation" of the gene encoding DPD that is associated with DPDD, and provides a novel method for the efficient detection of this DPDD splicing mutation.

Example 1

The PCR reaction is performed using human genomic DNA obtained from fibroblast cultures of 5 members of a family, one of which was afflicted with DPDD.

The PCR primers were selected so as to bracket the exon that is not present in the DPDD gene and part of the two introns on either side of this exon. Primers were synthesized using an Applied Biosystems 394 DNA & RNA synthesizer. The forward primer was DELF1, which encompassed nucleotide 154 to nucleotide 175 of the DPFD gene sequence and had the sequence TGCAAATATGTGAGGAGGGACC (*see also* Figure 1 and Figure 2). The reverse primer was DELR1, which encompassed nucleotide 563 to nucleotide 542 of the DPFD gene sequence and had the sequence CAAAGCAACTGGCAGATTC (*see also* Figure 1 and Figure 2). PCR was carried out in

50 μl of a reaction mixture consisting of 10 m*M* Tris-HCl pH 8.3, 50 m*M* KCl, 2.5 m*M* $MgCl_2$, 0.5 m*M* dNTPs, 1 μ*M* of each primer, and 2.5 units *Taq* polymerase (Roche Molecular Systems) for 30 cycles denaturing at 96°C for 1 min, annealing at 55°C for 1 min, and extending at 72°C for 2 min. The amplified products were extracted with 1 volume of chloroform and purified by filtration through Centricon 100 filter units.

Typically, one-fifth of the purified PCR product was used for sequence analyses with an Applied Biosystems 373A automated sequencer and fluorescent dye-deoxy terminator chemistry. Sequence data were analyzed using the MacVector sequence analysis software (International Biotechnology).

The PCR fragments were analyzed by electrophoresis in a 1% agarose gel stained with ethidium bromide. For Southern blot analysis, the DNA in the gels were depurinated by a 20-min incubation in 200 m*M* HCl and denatured by a 20-min incubation in 0.5 *M* NaOH. The DNA was transferred to Gene Screen Plus nylon membranes (New England Nuclear) overnight in 0.5 *M* NaOH, and was fixed to the filter by baking for 90 min at 80°C. Prehybridization was done in a solution of 6 x SSC, Denhardt's reagent, 0.5% $NaDodSO_4$, and 0.2 mg/ml sonicated salmon sperm DNA at 65°C for 3 hr, then hybridization was carried out overnight at 65°C in the same solution containing 1.5 z 10_6 cpm/ml of ^{32}P random priming-labeled human DPD, cDNA. After washing at 65°C for 20 min in 2 x SSC, 0.5% $NaDodSO_4$, and 45 min in 0.1 x SSC, 0.5% $NaDodSO_4$ at 65°C, the membranes were exposed to X-ray film (Eastman Kodak Co.) at -80°C for 30 min.

The results of sequencing the various genomic DNAs reveal the presence of a G to A mutation at position 434. This mutation was located within the splice site. The DPDD patient was homozygous for this mutation. Both parents and one sibling were heterozygous and another was homozygous for the normal allele.

The G to A mutation resulted in the destruction of a restriction site recognized by restriction endonucleases which recognize the Mae II binding sequence, ACGT. Consequently, after PCR amplification of genomic DNA using the primers

DELF1 and DELR1, digestion of the resulting PCR product with a restriction endonuclease that cleave the MaeII site yields fragments having the f ll wing sizes:

Normal allele	Mutant allele
278 + 131 bp	408 bp

Example 2: Identification of a Polymorphism which causes DPDD

Family pedigree

Blood samples from a British family were collected after a family member displayed excessive 5-FU toxicity during chemotherapy. This family is unrelated to any family with DPD deficiency that has previously been reported. To screen for the presence of the mutation leading to the deficiency in the normal population, genomic DNA was collected from healthy subjects representing the following ethnic groups: Caucasians (English and Finnish), Asians (Japanese and Taiwanese) and African-Americans.

DPD catalytic activity.

DPD catalytic activity for the family under study was determined from peripheral blood mononuclear cells using a previously described HPLC method (McMurrough *et al.* (1996) *Br. J. Clin. Pharmacol.*). In brief, peripheral blood mononuclear cells were purified from the patient's blood in a density gradient using Ficoll-Hypaque (Pharmacia, Upsala) and incubated for 90 min with [^{14}C]-5 FU. The production of 5-FU metabolites were then quantified by HPLC analysis by radiodetection.

Isolation of RNA and RT-PCR.

Total RNA was isolated from peripheral blood mononuclear cells by the guanidinium thiocyanate-phenol-chloroform method (Chomezynski *et al.* (1987) *Anal. Biochem.* 162: 156-159). The RNA solutions were dissolved in DEPC-treated water and stored at -80°C until use. RT-PCR was performed as described (Meinsma *et al.* (1995) *DNA Cell. Biol.* 14: 1-6) and the products were analyzed by electrophoresis in 0.8% agarose gels and visualized by staining with ethidium bromide.

Cloning and characterization of the genomic fragment encoding the deleted exon.

The DPD cDNA was used as a probe to isolate a P1 clone containing about 100 kbp of the human *DPD* gene (PAC 5945) from a high density PAC human genomic library (Genome Systems, St. Louis, MO). Southern blotting was used to confirm that the P1 clone contained the deleted exon using a probe located within the deleted fragment.

This probe was synthesized from the DPD cDNA by using the primers: DPD15 (forward): 5' TTGTGACAAATGTTTCCC 3' and DPD15R (reverse): 5' AGTCAGCCTTTAGTTCAGTGACAC 3' to specifically amplify the putative exon. PCR conditions were as indicated below but extension was carried out at 72°C for 1 min. This PCR fragment was purified using a Wizard PCR purification kit (Promega, Madison, WI), labeled with [α^{32}P]-dCTP and hybridized with the clone PAC 5945. DNA was purified from this genomic clone using Qiagen columns (Qiagen, Chatsworth, CA), and the 5' and 3' ends of the deleted exon and adjacent intronic regions were sequenced by chromosome walking from within the deleted exon using dideoxy terminator chemistry and an ABI 373A DNA sequencer (Applied Biosystems, Foster City, CA). The intronic sequences obtained allowed the selection of appropriate primers (*e.g.*, delF1 and delR1) to amplify from genomic DNA the complete exon and immediate flanking intronic sequences. All primers used in this study were synthesized with a 391 DNA&RNA synthesizer (Applied Biosystems, Foster City, CA).

Genotyping of the mutant *DPD* allele.

A 409 bp PCR genomic fragment corresponding to the deleted exon (from G1822 to C1986 in Yokota, *et al.* (1994) *J. Biol. Chem.* 269: 23192-23196) plus the flanking intronic sequences containing the AG and GT splicing consensus sequences was amplified from human genomic DNA using the primers delF1 (forward) 5' TGCAAATATGTGAGGAGGGACC 3' and delR1 (reverse) 5' CAGCAAAGCAACTGGCAGATTC 3'. PCR amplification was carried out in a 100 μl reaction mixture containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 0.5 μM of each primer, 2.5 units of *Taq* Polymerase (Roche Molecular Systems) and 500 ng of genomic DNA template for 31 cycles by denaturing at 94°C for 1 min, annealing at 60°C for 1 min and extending at 72°C for 2 min. Subjects identified as wild type, heterozygous or homozygous for the splicing mutation could be distinguished by digesting the PCR product with the restriction endonuclease Mae II (Boehringer, Mannheim, Indianapolis, IN) and electrophoresis in 1% regular, 3% NuSieve agarose gels (FMC Bioproducts, Rockland, ME). The genotypes obtained were verified by sequencing the 409 bp PCR product. The sources of the genomic DNA samples for the different ethnic groups correspond to those previously described (Fernandez-Salguero *et al.* (1995) *Am. J. Hum. Genet.* 57: 651-660).

DPD catalytic activity.

To determine DPD catalytic activities in the proband and family members, lymphocytes were isolated and subjected to analysis by HPLC (McMurrough *et al.* (1996) *Br. J. Clin. Pharmacol.*). Activities ranged between 24% to 70% of the mean DPD activity determined in the normal population (155 pmole/min/mg protein) (Table 1). It should be noted that all the subjects in the pedigree presented a relatively low catalytic activity (below 70% of the mean in the normal population). Subject I1, a cancer patient who exhibited 5-FU toxicity, had catalytic activity which was 24% of the mean population.

RT-PCR analysis.

RT-PCR was performed on total RNA isolated from peripheral blood mononuclear cells corresponding to each of the members of the pedigree shown in Figure 1. The full length coding region of the DPD cDNA was amplified in three fragments of 1.5 kbp, 906 bp and 919 bp. The 1.5 kbp and 906 bp PCR fragments were detected in all the members of the family and correspond to the normal mRNA as expected from the cDNA sequence. A 741 bp band corresponded to the deleted exon and was heterozygous with respect to the wild type 906 bp fragment. Sequencing of this 741 bp PCR fragment confirmed that a 165 bp deletion has occurred that precisely corresponds to that previously reported in a Dutch family (Meinsma *et al.* (1995) *DNA Cell. Biol.* 14: 1-6). No homozygous mutant subjects were found in this pedigree. The generation of the mutant mRNA is illustrated in Figure 3. Faulty splicing is the mechanism for the generation of the shorter mRNA, which due to the absence of 55 amino acids would be unable to be translated into an intact DPD enzyme.

Genomic DNA sequence analysis of the mutant *DPD* allele.

A clone containing the deleted exon was isolated in a recombinant P1 phage. Sequence of the intronic regions flanking the deleted exon enabled the development of a PCR reaction to sequence the exon and across the 5' and 3' splicing sites. Sequence of the 409 bp PCR genomic fragment clearly indicated the presence of a G to A point mutation (GT to AT) at the 5' donor splice consensus sequence in the British pedigree in subjects I1, II2, II3, II4, II7, II8, II9, III6, III8 and III9. Subjects II1, II6, III1, III2, III4 and III7 were found to have two normal alleles. The subject previously reported to have the 165 bp deletion present in both alleles and his heterozygote brother (subjects 1 and 2 in Meinsma *et al.* reference) were also analyzed and found to possess the splicing mutation

present in both alleles for subject 1 and in one allele for subject 2. The G to A mutati n correlates 100% with the corresponding RT-PCR results for the deleti n of the entire exon, indicating that the CT to AT mutation in the 5' splicing consensus sequence of the *DPD* gene leads to skipping of the entire preceding exon. The genotypes and catalytic activities for this 5'-splicing mutation in the pedigree are summarized in Table I.

Genotyping test for the mutant *DPD* allele and population screening.

The 409 bp genomic PCR fragment containing the deleted exon has a restriction site for the endonuclease Mae II (A-CGT) which produces two fragments of 278 bp and 131 bp from the wild type allele. This site is eliminated when a G to A mutation occurs at the 5'-splicing site (A-CAT). A diagrammatic representation of the strategy developed for PCR amplification of the 409 bp fragment and the restriction enzyme pattern expected with MAE II is shown in Figure 4. The reliability of the genotyping assay was confirmed by sequencing the 409 bp PCR product in these samples. By analyzing normal subjects within different ethnic groups (Table II), we found that this mutant allele is present in the Finnish (4.4%) and Taiwanese (5.4%) populations where heterozygotes were detected. However, within the samples tested to date, no mutant alleles were found in 35 Japanese or 20 African-American subjects. All heterozygous subjects found by restriction enzyme analysis were also confirmed by genomic DNA sequence analysis.

TABLE I: *Relationship between phenotype and genotype for the DPD catalytic activity and the presence of the 5'-splicing mutation in the pedigree analyzed.*

Family Member	DPD activity (%)*	Genotype#
I1	24	+/-
II1	57	+/+
II2	39	+/-
II3	36	+/-
II4	42	+/-
II5	n/d	n/d
II6	35	+/+
II7	58	+/-
II8	22	+/-
II9	28	+/-
III1	55	+/+
III2	60	+/+
III3	n/d	n/d
III4	60	+/+
III5	n/d	n/d
III6	36	+/-
III7	70	+/+
III8	31	+/-
III9	56	+/-

* DPD activity is expressed as percentage of the mean activity in the normal population.

Genotype for the GT to AT splice mutation: +/+: wild type; +/-: heterozygote; /: homozygote; n/d: not determined

TABLE II: *Allelic frequency for the presence of the 5'-splicing mutation in different ethnic groups from the normal population*

	Subject*	Wild type (+/+)	Heterozygote (+/-)	Homozygote (-/-)
Caucasian				
British	30	100 (30)[a]	0 (0)[a]	0 (0)[a]
Finnish	45	95.6 (43)	4.4 (2)	0 (0)
Asian				
Taiwanese	37	94.6 (35)	5.4 (2)	0 (0)
Japanese	35	100 (35)	0 (0)	0 (0)
African-American	20	100 (20)	0 (0)	0 (0)

*The allelic frequency is indicated as a percentage of the total number of subjects analyzed. The absolute number of heterozygote subjects are indicated in parenthesis.

All publications, patents and patent applications mentioned in this specification are herein incorporated by reference into the specification to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated herein by reference.

Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claims.

WHAT IS CLAIMED IS:

1. A method of detecting a splicing defect in a dihydropyrimidine dehydrogenase gene, comprising determining whether a genomic DNA encoding the dihydropyrimidine dehydrogenase gene has a wild-type intron-exon boundary for an exon which encodes amino acids 581-635 of a corresponding wild-type dihydropyrimidine dehydrogenase protein.

2. The method of claim 1, wherein the method comprises the step of amplifying intronic genomic DNA encoding the dihydropyrimidine dehydrogenase in the region flanking the exon which encodes amino acids 581-635.

3. The method of claim 2, wherein the method comprises amplifying the genomic DNA with a primer which hybridizes to a dihydropyrimidine dehydrogenase intronic nucleic acid which hybridizes to a primer selected from the group of primers consisting of DELF1, and DELR1 under stringent conditions.

4. The method of claim 2, wherein DNA amplified with the primers is cleaved with a restriction endonuclease which recognizes a Mae II cleavage site.

5. The method of claim 1, wherein the sequence of the intron-exon boundary is determined using an oligonucleotide array.

6. A method of screening patients for sensitivity to 5-fluorouricil, comprising isolating a genomic DNA from the patient which encodes the dihydropyrimidine dehydrogenase gene and determining whether the gene has a wild-type intron-exon boundary for an exon which encodes amino acids 581-635 of a corresponding wild-type dihydropyrimidine dehydrogenase protein.

7. The method of claim 6, wherein the method comprises the step of amplifying intronic genomic DNA encoding the dihydropyrimidine dehydrogenase in the region flanking the exon which encodes amino acids 581-635.

8. The method of claim 7, wherein the method comprises amplifying the genomic DNA with a primer which hybridize to a dihydropyrimidine dehydrogenase intronic nucleic acid which hybridizes to a primer selected from the group of primers consisting of DELF1 or DELR1 under stringent conditions.

9. The method of claim 7, wherein DNA amplified with the primers is cleaved with a restriction endonuclease which recognizes a Mae II cleavage site.

10. A composition comprising a first PCR primer which binds to DNA 3' of a splice site in the genomic DNA for dihydropyrimidine dehydrogenase gene for an exon encoding amino acids 581-635, and a second PCR primer which binds to DNA 5' of a splice site in the genomic DNA for dihydropyrimidine dehydrogenase gene for an exon encoding amino acids 581-635.

11. The composition of claim 10, wherein the first PCR primer binds to intronic dihydropyrimidine dehydrogenase DNA.

12. The composition of claim 10, wherein the second PCR primer binds to intronic dihydropyrimidine dehydrogenase DNA.

13. The composition of claim 10, wherein the first PCR primer hybridizes under stringent conditions to a nucleic acid complementary to DELF1.

14. The composition of claim 10, wherein the second PCR primer hybridizes under stringent conditions to a nucleic acid complementary to DELR1.

15. A kit comprising a container, a first PCR primer which binds to DNA 3' of a splice site in the genomic DNA for dihydropyrimidine dehydrogenase gene for an exon encoding amino acids 581-635, and a second PCR primer which binds to DNA 5' of a splice site in the genomic DNA for dihydropyrimidine dehydrogenase gene for an exon encoding amino acids 581-635.

16. The kit of claim 15, wherein the kit further comprises instructions for the detection of splicing site defects in the dihydropyrimidine dehydrogenase gene.

17. The kit of claim 15, wherein the kit further comprises Mae II.

18. The composition of claim 15, wherein the first PCR primer hybridizes under stringent conditions to a nucleic acid complementary to DELF1.

19. The composition of claim 15, wherein the second PCR primer hybridizes under stringent conditions to a nucleic acid complementary to DELR1.

```
1                                                                    50
TGTTAATGAA GATAAATATT TTTGTTTTTT CGCTGTTCTA AACCTAGGGT
ACAATTACTT CTATTTATAA AAACAAAAAA GCGACAAGAT TTGGATCCCA

51                                                                  100
TACAAGAAGT AATTTATCTG GAGCTAACAA ATACTTTATT TTACCTTTTT
ATGTTCTTCA TTAAATAGAC CTCGATTGTT TATGAAATAA AATGGAAAAA

101                                                                 150
ATTTGCAAGT AGTTTATGTT CAATTCTAAT TTAATGTATA TTAAAAATTC
TAAACGTTCA TCAAATACAA GTTAAGATTA AATTACATAT AATTTTTAAG

151                                                                 200
  5′  PRIMER DELF1          3′
CTCTGCAAAT ATGTGAGGAG GGACCTCATA AAATATTGTC ATATGGAAAT
GAGACGTTTA TACACTCCTC CCTGGAGTAT TTTATAACAG TATACCTTTA

201                                                                 250
GAGCAGATAA TAAAGATTAT AGCTTTTCTT TGTCAAAAGG AGACTCAATA
CTCGTCTATT ATTTCTAATA TCGAAAAGAA ACAGTTTTCC TCTGAGTTAT

251                                                                 295
TCTTTACTCT TTCATGAG GAC ATT GTG ACA AAT GTT TCC CCC AGA
AGAAATGAGA AAGTAGTC CTG TAA CAC TGT TTA CAA AGG GGG TCT
                  D   I   V   T   N   V   S   P   R
                      5′                      3′

296                                                                 337
ATC ATC CGG GGA ACC ACC TCT GGC CCC ATG TAT GGC CCT GGA
TAG TAG GCC CCT TGG TGG AGA CCG GGG TAC ATA CCG GGA CCT
 I   I   R   G   T   T   S   G   P   M   Y   G   P   G

338                                                                 379
CAA AGC TCC TTT CTG AAT ATT GAG CTC ATC AGT GAG AAA ACG
GTT TCG AGG AAA GAC TTA TAA CTC GAG TAG TCA CTC TTT TGC
 Q   S   S   F   L   N   I   E   L   I   S   E   K   T

380                                                                 421
GCT GCA TAT TGG TGT CAA AGT GTC ACT GAA CTA AAG GCT GAC
CGA CGT ATA ACC ACA GTT TCA CAG TGA CTT GAT TTC CGA CTG
 A   A   Y   W   C   Q   S   V   T   E   L   K   A   D
                         3′                          5′
```

FIG. 1A.

Mae II

```
422                                                          470
TTC CCA GAC AAC G TAAGTG TGATAAAAAT CTAAAACAAG AGAATTGGCA
AAG GGT CTG TTG C ATTCAC ACTATTTTTA GATTTTGTTC TCTTAACCGT
 F   P   D   N   X

471                                                          520
TAAGTTGGTG AATGTTTATT TAAACATCCA ATTCATAGGC TTATAAATAT
ATTCAACCAC TTACAAATAA ATTTGTAGGT TAAGTATCCG AATATTTATA

521                                                          570
TAATGTGTAT ATTTTATCAA CGAATCTGCC AGTTGCTTTG CTGATGCATA
ATTACACATA TAAAATAGTT GCTTAGACGG TCAACGAAAC GACTACGTAT
                       3' PRIMER DELR1          5'
571                                                          620
GAAAGATAAA AAAGAAAGAA AAGCTCAAGA ACTCATAAAA ACCCACACAA
CTTTCTATTT TTTCTTTCTT TTCGAGTTCT TGAGTATTTT TGGGTGTGTT

621                                                          670
TGTGAAGCTC TGTTATAAAT GGGTGCCATG TAAGATGGAA GAAGTATCTA
ACACTTCGAG ACAATATTTA CCCACGGTAC ATTCTACCTT CTTCATAGAT

671                                                          720
CATAAGCAGA AGGAAGAGAA ATGAAATACT CATTTTATTG AGTTGGCCCC
GTATTCGTCT TCCTTCTCTT TACTTTATGA GTAAAATAAC TCAACCGGGG

721                                                          770
CACTGTATGT GGCTGGTATT TATGAAGGTG ATGACCCAGG AAGAAATTGT
GTGACATACA CCGACCATAA ATACTTCCAC TACTGGGTCC TTCTTTAACA

771                                                          820
AAACTATAAA CCACTCCAAA TATAAACCCG AGGCAGAAGC AGCATATCTC
TTTGATATTT GGTGAGGTTT ATATTTGGGC TCCGTCTTCG TCGTATAGAG

821                                                  861
CTATGAAGCC TGTATTTACT CAGTGGGAAA TAATTTATTA A
GATACTTCGG ACATAAATGA GTCACCCTTT ATTAAATAAT T
```

FIG. 1B.

5′ TGTTAATGAAGATAAATATTTTTATTTTTGCGCTATTCTAAACCTAGAGTTAC

AAGAAGTAATTTATCTGGAGCTAACAAATACTTTATTTTACCTTTTTATTTGCAA

5′ TGCAAATATG

GTAGTTTATGTTCAATTCTAATTTAATGTATATTAAAAATTCCTCTGCAAATATG

TGAGGAGGGACC 3′

TGAGGAGGGACCTCATAAAATATTGTCATATGGAGATGAGCAGATAATAAAGA

TTATAGCTTTTCTTTGTCAAAAGGAGACTCAATATCTTTACTCTTTCATCAGGA

CATTGTGACAAATGTTTCCCCCAGAATCATCCGGGGAACCACCTCTGGCCC

CATGTATGGCCCTGGACAAAGCTCCTTTCTGAATATTGAGCTCATCAGTGAG

AAAACGGCTGCATATTGGTGTCAAAGTGTCACTGAACTAAAGGCTGACTTCC

CAGACAACGT AAGTGTGATAAAAATCTAAAACAAGAGAATTGGCATAAGTT

GGTGAATGTTTATTTAAACATCCAATTCATAGGCTTATAAATATTAATGTGTATA

3′ CTTAGACGGTCAACGAAACGAC 5′

TTTTATCAAAGAATCTGCCAGTTGCTTTGCTGATGCATAGAAAGATAAAAAAG

AAAGAAAAGCTCAAGAACTCATAAAAACCCACACAATGTGAAGCTCGTTATA

AATGGGTGCCATGTAAGATGGAAGAAGTATCTACATAAGCAGAAGGAAGAGA

AATGAAATACTCAATTTATTGAGTTGGCCCCCACTGTATGTGGCTGGCATTTA

TGAAGGTGATGACCCAGGAAGAAATTGTCACCTATAAATCATCCAAATATCC

CGAGGCAGAAGCAGCATCTCTCCTATGAAGTCTGTATTTATTTCAGCGGGAA

ATAATTTATTA 3′

FIG. 2.

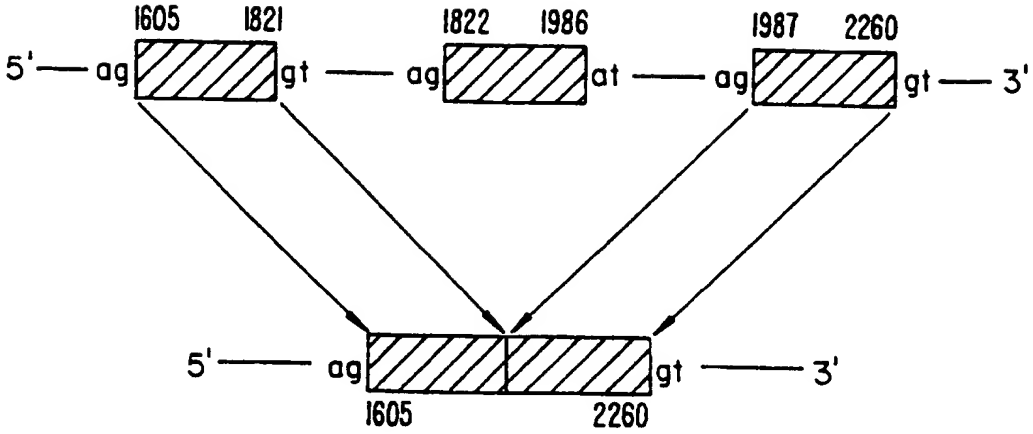
1605
1821
1822
1986
1987
2260
5'
ag
gt
ag
at
ag
gt
3'
5'
ag
gt
3'
1605
2260

FIG. 3.

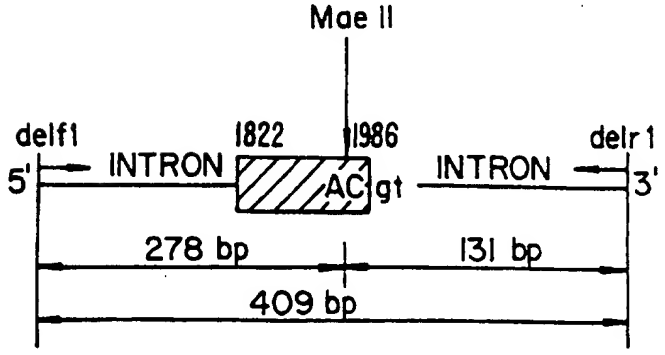
Mae II
delf1
1822
1986
delr1
5'
INTRON
AC
gt
INTRON
3'
278 bp
131 bp
409 bp

FIG. 4.

INTERNATIONAL SEARCH REPORT

In. ..ational Application No
PCT/US 97/04269

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	DNA AND CELL BIOLOGY, vol. 14, no. 1, January 1995, page 1-6 XP000675568 MEINSMA ET AL: "Human polymorphism in drug metabolism: mutation in the dihydropyrimidine dehydrogenase gene results in exon skipping and thymine uracilurea" cited in the application see abstract and paragraph 1, page 5. --- -/--	1-19

[X] Further documents are listed in the continuation of box C. [ ] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
11 June 1997	07.07.97

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+ 31-70) 340-3016	Osborne, H

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 97/04269

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	TRENDS PHARMACOL. SCI., vol. 16, no. 10, October 1995, pages 325-7, XP000674605 GONZALEZ, F ET AL: "Diagnostic analysis, clinical importance and molecular basis of dihydropyrimidine dehydrogenase deficiency" see page 326, paragraph 2 - page 327, paragraph 3	1-19
Y	SINGAPORE JOURNAL OF OBSTETRICS AND GYNACOLOGY, vol. 26, no. 3, November 1995, pages 176-86, XP000600337 ROY, A ET AL: "molecular scanning of human diseases" see the whole document	1-19
Y	NUCLEIC ACIDS RESEARCH, vol. 15, no. 14, 1987, pages 5613-28, XP002032865 MARVIT, J ET AL: "GT to AT transition at a splice donor site causes skipping of the preceeding exon in phenylketonuria " see abstract	1,6
Y	JOURNAL BIOLOGICAL CHEMISTRY, vol. 265, no. 20, July 1990, pages 12067-74, XP002032866 KUIVANIEMI, H. ET AL: "Identical G to a mutations in three different introns of the type III procollagen gene (COL3A1) produce different patterns of RNA splicing in three variants of Ehlers-Danlos Syndrome IV" see abstract	1,6
P,X	JOURNAL OF CLINICAL INVESTIGATION, vol. 98, no. 3, August 1996, pages 610-15, XP000675565 WEI, X. ET AL: "Molecular basis of the human dihydropyrimidine dehydrogenase deficiency and 5-fluorouracil toxicity" see the whole document	1-19

1